DOCUMENT RESUME

ED 227 441 CS 007 009

AUTHOR	Pearson, P. David, Ed.; Hansen, Jane, Ed.
TITLE	Reading: Disciplined Inquiry in Process and Practice. Twenty-Seventh Yearbook of the National Reading Conference
INSTITUTION	National Reading Conference, Inc.
PUB DATE	78
NOTE	302p.; Proceedings of the annual meeting of the National Reading Conference (27th, New Orleans, LA, December 1-3, 1977).
PUB TYPE	Collected Works - Conference Proceedings (021) -- Reports - Research/Technical (143) -- Information Analyses (070)
EDRS PRICE	MF01/PC13 Plus Postage.
DESCRIPTORS	Adults; *Cognitive Processes; Elementary Secondary Education; Evaluation Methods; Higher Education; Language Processing; Program Development; Reading Attitudes; *Reading Comprehension; Reading Diagnosis; *Reading Instruction; *Reading Processes; *Reading Programs; *Reading Research; Study Skills; Teacher Education; Word Recognition
IDENTIFIERS	National Reading Conference (Organization)

ABSTRACT

The papers in this collection represent a wide spectrum of approaches, philosophies, viewpoints, and techniques of scholarly endeavor in their treatment of reading and its cognitive processes. The 48 articles, presented at the 1977 meeting of the National Reading Conference, are arranged according to the following categories: (1) reading and teacher education, (2) reading and program development, (3) reading and word identification, (4) reading and comprehension, (5) reading and language, (6) reading and assessment, (7) reading and the aged, (8) reading and the affective domain, (9) reading and research, and (10) reading and study skills. The collection concludes with a review of 1975-76 research on college and adult reading, and the program from the 1977 National Reading Conference. (HTH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED227441

READING: DISCIPLINED INQUIRY IN PROCESS AND PRACTICE

TWENTY-SEVENTH YEARBOOK

OF

THE NATIONAL READING CONFERENCE

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Edited by

P. DAVID PEARSON

University of Minnesota

JANE HANSEN

University of Minnesota

Published by

The National Reading Conference, Inc.
Clemson, South Carolina 29631

1978

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

The National Reading Conference, Inc.

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS 007009

Copyright, 1978
THE NATIONAL READING CONFERENCE, INC.
All rights reserved. This book or parts thereof
may not be reproduced in any form without
permission of the publishers

Distributed by
THE NATIONAL READING CONFERENCE
Godfrey Hall
Clemson University
Clemson, South Carolina 29631

The ideas expressed in the papers published
in the year book are those of the authors and
do not necessarily represent the viewpoints
of the National Reading Conference

PRINTED IN THE UNITED STATES OF AMERICA
BY
CONTINENTAL GRAPHICS, INCORPORATED
St. Paul, Minnesota

OFFICERS AND BOARD OF DIRECTORS
OF THE
NATIONAL READING CONFERENCE, INC.

J. JAAP TUINMAN, *President*
Simon Fraser University

HARRY SINGER, *Vice-President*
University of California

GORDON GRAY, *Treasurer*
Clemson University

VONCILE C. MALLORY, *Secretary*
Florida Atlantic University

EDWARD FRY, *Past-President*
Rutgers University

FRANK GREENE, *Member*
McGill University

JOANNA WILLIAMS, *Member*
Columbia University

GARR CRANNEY, *Member*
University of Florida

S. JAY SAMUELS, *Member*
University of Minnesota

RICHARD ROBINSON, *Convention Manager*
University of Missouri

P. DAVID PEARSON, *Yearbook Editor*
University of Minnesota

J. JAAP TUINMAN, *Editor*
Journal of Reading Behavior
Simon Fraser University

EDITORIAL ADVISORY COMMITTEE
27th NRC YEARBOOK

DICK ALLINGTON
State University of New York at Albany

EUNICE ASKOV
Penn State University

EMERY BLIESMER
Penn State University

RICHARD BLOOMER
University of Connecticut

DON CUNNINGHAM
Indiana University

JAMES CUNNINGHAM
University of North Carolina - Chapel Hill

PATRICIA CUNNINGHAM
Alamance County Schools - Graham, NC

BETH DAVEY
University of Maryland

PRISCILLA DRUM
University of California - Santa Barbara

KEN DULIN
University of Wisconsin

FRANK GREENE
McGill University

DOROTHY HANSEN
Brigham Young University

JEROME C. HARSTE
Indiana University

AL KINGSTON
University of Georgia

MICHAEL KAMIL
Purdue University

RON LESLIE
New York University

MARTHA MAXWELL
University of California - Berkeley

BONNIE J. F. MEYER
Arizona State University

BEVERLY MORRISON
University of Wisconsin

CHRIS RAMIG
Georgia State University

WILLIAM RUPLEY
Texas A&M University

MARGARET SMITH-BURKE
New York University

ELIZABETH SULZBY
Northwestern University

ROBERT TIERNEY
University of Arizona

JOSEPH VAUGHAN
University of Arizona

DAVID M. WARK
University of Minnesota

TABLE OF CONTENTS

Page

READING AND TEACHER EDUCATION

Preservice teachers' perceptions of reading instruction 1
Noel Bowling, Joseph Muia, and Gary Shaffer

An effective inservice model for content area reading in the secondary schools 6
Eunice Askov, Mary Dupuis, and Joyce Lee

The effectiveness of informal assessment questions constructed by secondary teachers 13
Carol Davis

Observed patterns of teacher-pupil classroom behavior as predictors of student growth in reading 16
Jeffrey L. Lorentz and Homer Coker

Teacher inferences of the characteristics of non-standard speaking readers 20
Christopher Ramig and Robert C. Granger

READING AND PROGRAM DEVELOPMENT

The effects of differing materials on the reading process 27
John Stansell, Jerome C. Harste, and Roger J. DeSanti

Effects of impulsity-reflectivity and type of phonics instruction on reading achievement 36
John Readence and R. Scott Baldwin

READING AND WORD IDENTIFICATION

The development of orthographic sensitivity during the school year by primary grade children 41
Jerome Niles and Barbara Taylor

Experiments in word learning 45
Elizabeth Arnold, George McNinch, and Wallace Miller

Children's explanations of word similarities in relation to word knowness 51
Elizabeth Sulzby

Word prediction of good and poor readers 56
Richard Allington and Michael Strange

READING AND COMPREHENSION

The acquisition of knowledge from text 61
Gregory Pearson

Inferential operations of children involved in discourse processing 68
C. Bridge, R. Tierney, and M. Cera

Page

Comprehension profiles of good and poor readers across materials of increasing difficulty 73
Jill Olshavsky

The relationship between reading ability and semantic verification tasks 77
Karen Hanson, Robert Schreiner, and Thomas Hummel

Perceptual vs. semantic information processing in semantic category decisions 84
Michael Kamil and Raymond Hanson

Prose recall responses and categories for scoring 88
Priscilla Drum

Two factors affecting text recall 93
Roger Lantaff

A system for scoring readers' recall of propositions from texts and cumulative effects of information on inference making 99
James Flood

What makes reading difficult: the complexity of structures 106
Eunice Schmidt

Immediate memory for sentences of fast and slower readers as a function of rate of presentation 111
Nancy Nelson and John Geyer

Investigating the "print to meaning" hypothesis 116
Patricia Cunningham and James Cunningham

A partial validation of the kernel distance theory for readability 121
Edward Fry, Jane Weber, and Joseph DePierro

Fourth graders' comprehension of story structures under three recall conditions 125
Laura Chodos and Peter Mosenthal

Differential effects of prior context, style and deletion patterns on cloze comprehension 133
Margaret Smith-Burke, Patricia Gingrich, and Daniel Eagleeye

READING AND LANGUAGE

Some things the reader needs to know that the listener doesn't 138
Glenn Kleiman and Diane L. Schallert

A taxonomy of language experiences 143
Andee Rubin

Page

Characteristics of the cloze procedure as a research tool in the study of language 148
Earl Rankin

The relation of comprehension to semantic and syntactic language cues utilized during oral and silent reading 154
R. Ann Zinck

READING AND ASSESSMENT

Assessing reading performance at the secondary level through the utilization of a cognitive self-rating scale 161
Charles Peters

The comprehension of prose: problems in measuring learning outcomes 166
J. Bleakley and E. Johnson

Reliability of the cloze procedure as assessments of various language elements 175
Joseph Vaughan and Keith Meredith

Stability of cloze scores across varying deletion patterns 181
Keith Meredith and Joseph Vaughan

The reliability of the maze procedure for intermediate and junior high school students 185
John Bradley and Keith Meredith

The use of an unobtrusive screening device to approximate adult reading levels 190
Ellen West

READING AND THE AGED

Future shock and the aged: is reading a cure or part of the problem? 193
Molly Wilson

Problems in the measurement of intellectual and linguistic capacities, including reading, in the aged 197
Terry Lovelace

READING AND THE AFFECTIVE DOMAIN

Self-perceptions and reading habits of adolescents, college students, and adults 202
Patricia Anders and Douglas Cardell

READING AND RESEARCH

Stimulus control in the study of reading 206
George McConkie

Page

Some reasons for focusing on classrooms in reading research 213
Ray McDermott

READING AND STUDY SKILLS

Influence of a structured overview on comprehension and oral reading miscues of selected college students 218
Thomas Bean

The effect of expanded directions and adjunct aids on students' comprehension of world history text 222
Richard Vacca

Effects of sustained silent reading and reading skills instruction on changes in secondary students reading attitudes and achievement 226
Anne Wolf

An investigation of the REAP reading/study procedure: its rationale and efficacy 229
Marilyn Eanet

The effect of creative thinking-reading activities (CT-RA) on reading comprehension 233
Martha Haggard

The effects of instruction in testwiseness in a college reading improvement course 237
E. Jongsma, Ronald E. Pound, and Martha L. Tips

Comparison of the effectiveness of three study techniques for college students 242
Vincent Orlando and Dristina Hayward

The competencies that junior college chairpersons expect from their reading course graduates 243
Helen Covington and Lee Mountain

1975-1976 review of research on college-adult reading 250
Edward Summers, Anne Forester, and Sharon Jeroski

The Program from 1977 Conference 286

FOREWARD

This volume represents the disciplined inquiry into the concerns, questions, and hypotheses of many persons involved in the world of reading. Some articles pose questions, some seek to answer questions, and others may even have found a few answers.

A particular message of appreciation goes to the editorial board for weighing the strengths and weaknesses of the various manuscripts. Their cooperation deserves an explicit note of thanks.

In addition, several persons at the University have assisted in the process. A special thank you goes to Karen Ritz for designing our cover. Georganna Sampson has found herself in the middle of things frequently. Irma Chinn, Judy Tebbitt, and Josephine Zimmer's secretarial pool have all typed and mailed many papers.

It has been interesting to read (and reread) all of the articles. Hopefully, the NRC membership will find them as interesting, informative and helpful as we did. If they stimulate further research that sheds light on our concerns about processes and practices in reading, the volume has served a worthwhile purpose.

P. David Pearson
Jane Hansen
University of Minnesota

NOEL BOWLING
West Virginia College of Graduate Studies

JAMES LAFFEY
JOSEPH MUIA
GARY SHAFFER
James Madison University

Preservice teachers' perceptions of reading instruction

"Teachers teach as they are taught and not as they have been instructed to teach," is a familiar hypothesis expressed by disheartened teacher educators. In spite of the wide-spread voicing of this and similar statements, a survey of the voluminous research literature respective to factors influencing teaching failed to yield a single study that has tested this proposition (Balzer, Evans & Blosser, 1973; Smith, 1969; Rosenshine & Furst, 1971). It is surmised by these writers that the failure to submit this hypothesis to analysis is due to the difficulty, perhaps impossibility, of obtaining data to substantiate the characteristics of prior instruction. The limiting effects of the time alone, specifically the gathering of data over a fourteen-or-more-year period, would tend to prohibit an investigation of this nature.

This study was initiated as a result of observed differences noted by the writers between preservice teachers' actual performances while conducting diagnostic teaching of reading instruction in public school settings and the suggested performances resulting from the instructional and resource information contained within a course of study entitled "Diagnostic Teaching of Reading." The purpose of this study was to design an instrument that would ascertain preservice teachers' perceptions of specific reading instructional practices that were being implemented in the classrooms during the periods when they received reading instruction. It was reasoned that if such an instrument could be developed, data concerning the influences of prior instruction upon the way one teaches could be obtained.

In other words, what preservice teachers perceived or believed to have been the instructional practices during the time when they experienced reading instruction was assumed by the authors to have a similar effect on the preservice teachers' perceptions of specific instruction practices were known by methods instructors prior to instruction related to these practices, then the instructor could adjust instruction accordingly.

Method

Subjects

At the start of the 1975 Spring semester, 110 preservice teachers, enrolled in seven sections of an advanced undergraduate reading course at James Madison University, responded to an instrument entitled *Preservice Teacher Perception Questionaire* (PTPQ). The mostly female subjects were second semester juniors who previously completed a reading course which contained a survey of instructional practices in reading. But for several possible exceptions, information from student data sheets indicated that the subjects' opportunities for formulating perceptions respective to instructional practices in reading were limited to vicarious and simulated experiences resulting from the previous reading class and those experiences received while the subjects were recipients of reading instruction.

Materials and Procedure

The PTPQ, designed to determine preservice teachers' perceptions of reading practices, consisted of an 83-item, self-report questionnaire. The items were designed by the authors and were partially based on the differences noted between suggested practices and actual performances observed of prior students in the practicum sessions of the course, "Diagnostic Teaching of Reading."

The format of the questionnaire called for two written responses per item. After reading the item the students were directed to first mark a space denoting the frequency an instructional reading practice was judged to have taken place. Always, sometimes, seldom, and never were response choices provided. In addition, the subjects were given the opportunity to mark outside of the frequency response box if they were unable to recall the instructional practice. This was labeled an "I forget" response. The second written response called for the subjects to judge the explicitness of the item. Clear and unclear choices were provided for this response.

In addition to the oral directions given to subjects for marking the items of the questionnaire, overall set for responding to the instrument was provided. The preservice teachers were instructed to try and remember their experiences when receiving reading instruction and to make their responses based on these memories.

Three forms of the questionnaire were administered simultaneously to each of the seven sections of preservice teachers. This was managed by having approximately a third of the subjects in each section respond to one of three forms. The three forms resulted from reordering of the 83 items.

Responses obtained from administering the PTPQ were analyzed in order to check the soundness of the instrument as a diagnostic device, as well as for interpreting purposes.

The estimates of soundness were based upon: (1) the explicitness of the items, (2) the consistency of the instrument, (3) the logical relationships existing between similar items, and (4) the variability of the responses. Interpretation of the data resulted from classifying related teaching practices into categories and then rank-ordering each practice according to the relative frequency indicated by the preservice teacher responses. In addition, patterns were interpreted by clustering conceptually-related and frequently-related items. Phi coefficient, chi^2, means, and standard deviation statistics were computed for each item. In determining the phi coefficient and chi^2 statistic, frequency counts of the response categories (always, sometimes, seldom and never) were combined to represent frequent and infrequent practices. The response categories "always" and "sometimes" formed the frequent practices, where as the "seldom" and "never" response formed the infrequent practices. The "I forget" response was not considered in the computation. A five-point scale was used for determining the mean and standard deviation for each item. All response categories were included in the scale with the "always" response considered as a five ranging to the "I forget" response as a one.

Results

The indicators of soundness of the PTPQ were found to be generally supportive. All 83 items were judged to be explicit on the basis that less than 5% of the total responses indicated that any one item was unclear. However the unclear responses did differentiate subjects. Seventeen of the 110 subjects accounted for more than 95% of the unclear responses, with one preservice teacher marking 32 of the 83 items "unclear."

Based upon an item-by-item, visual inspection of three frequency distributions resulting from the administration of the three forms, the PTPQ appears to be consistent. Frequency distributions for only three of the 83 items appeared to contain noticeable differences. However, when the adjacent response categories were combined into frequent and infrequent response groupings, the frequency distribution of each of the three forms was within a five (5) percentage-point range of the total frequency distribution for the item.

It was reasoned by the writers that (1) conceptually related teaching practices would be perceived by preservice teachers as occurring with similar frequency and (2) that conceptually opposite teaching practices would be perceived as occurring with reciprocal frequencies. Data supported the first proposition in that several clusters of similar teaching practices were perceived by the subjects to be occurring with similar frequency. For example, the four items ranked highest indicate a concern for a perfected oral reading performance. Generally the clustering of conceptually related practices by frequency was indicated throughout the 83-itemed PTPQ. On the other hand, the second proposition appears questionable. Whereas reciprocal statement pairs such as items 5 and 6 indicate a logically consistent relationship, items 17 and 18 are inconsistent.

Table 1

Perceptions of Teaching Practices Related to Oral Reading

Rank	Item[a]	Phi[b]	Chi^{2}[c]	*SD*	*M*	Concepts
1	29	.563	(.01)	.5	4.65	Read with expression
2	31	.506	(.01)	.9	4.38	Read accurately
3	28	.483	(.01)	.8	4.46	Read carefully
4	30	.470	(.01)	.8	4.34	Pronounce distinctly
5	17	.412	(.01)	.8	4.17	Called on students systematically
6	6	.348	(.01)	.9	3.95	Without prior preparation
7	45	.280	(.01)	1.0	3.56	Reread relevant segments
8	8	.235	(.01)	.8	3.61	Peer prompting
9	18	.166	(.05)	.8	3.65	Called on students randomly
10	24	.147	(.05)	1.2	3.35	Corrected for guessing
11	27	.125		1.3	3.35	Corrected for substitution-sound similarity
12	25	.098		1.1	3.35	Corrected for failure to attempt a word
13	19	.052		.9	3.39	Called on volunteers
14	44	.042		1.1	3.21	Reread entire story
15	26	.021		1.2	3.21	Corrected for semantic substitution
16	20	.005		1.0	3.29	Selected best reader for majority of reading
17	46	-.096		1.1	2.96	Clear rationale for rereading
18	47	-.128		1.1	2.66	Reread pts. of story w/difficult words
19	7	-.146	(.05)	1.0	2.94	If error, peer assigned to continue
20	5	-.191	(.01)	1.0	2.95	Prepare for oral reading

Note.

[a] 15% or more of the subjects could not remember this practice.

[b] Minus sign indicates that the majority of the preservice teachers perceived these practices as occuring infrequently.

[c] Numbers in parentheses indicate level of significance reached by applying Chi^{2} statistic.

It was further assumed by the authors that, if the PTPQ was to be of diagnostic use, it should contain items that addressed perceived teaching practices that range from very frequent to those that rarely occur. Keeping in mind that a 5-point scale was used in which the low frequency responses were valued at two for never and one for "I forget," there were eleven items with means of 4.00 and greater, 48 items with means between 3.00 and 3.99, and 24 items with means between 2.00 and 2.99. Variability of the items ranged between standard deviations of .5 to 1.5, with standard deviations for fifty-five of the 83 items being 1.0 and greater. Thus the instrument was judged to be effective in reflecting the individual differences of adult students.

In addition to analyzing the results for soundness purposes, the authors examined the data for indicators of the types of perceptions preservice teachers held respective to reading instructional practices. The 83 items of the PTPQ were classified into six categories – (1) oral reading, (2) word recognition, (3) comprehension, (4) classroom management, (5) general instructional practices, and (6) teaching practices related to ideas about reading. Following the classification process, the items in each category were rank ordered by frequency. This was done on the assumption that those behaviors perceived to be the more frequent would have greater impact on preservice teachers as a group than those teacher behaviors that were perceived to occur less frequently. Isolated teacher practices of significance occurred in each of the six categories; however, the results reported here will be limited to those trends interpretable from clusters of perceived teaching practices.

Seven general trends were suggested by the data. The first trend was formed from nine items that related to the teacher communicating expectations or criteria for student performance. This trend was perceived by preservice teachers to be the most frequently occurring of those determined by the instrument. Items 28, 29, 30 and 31 in Table 1 exemplify the perceived teaching behaviors that comprised this trend.

A second trend noted from the responses concerned the management of students' performances. The 22 teaching behaviors comprising this trend included determining of specific reading activities, grouping practices, and procedures for eliciting responses. While some of these practices vary in frequency from very high to very low, the majority of the practices were perceived by preservice teachers to be of frequent occurence. The third trend, was also perceived to be of frequent occurence. This trend was formulated by teaching practices related to correcting inappropriate student responses. Items in Table 1 such as 8, 24, 25, 26 exemplify the types of teaching behaviors that comprised this trend. In general the clusters of teaching behaviors that were perceived to occur frequently – trends one, two, and three, are at least tangentially related to the practical, how-to-do-it aspects of reading instruction.

The fourth, fifth, and sixth trends interpretable from the data were perceived by preservice teachers to occur infrequently. The fourth trend was formed from seven teaching practices related to preparing students for subsequent reading activities. The cluster includes items such as 6, 46, and 5 in Table 1. Trend four is logically related to trends one and three. An example of this relationship can be illustrated by the concepts reflected by items 5, 29, and 24 listed in Table 1. For example, if students were provided with opportunities to prepare for oral reading, then it seems likely that they would read expressively and, in addition, the teacher would not need to make frequent corrections for guessing. However, the data, suggests that this relationship did not hold in that teachers were remembered as seldom preparing students regarding the rationale for a task.

Trends five and six are more tentative in that there were only two or three items related to each cluster. Most preservice teachers perceived their reading teachers as only infrequently individualizing instruction (trend five) and using diagnostic precedures (trend six).

The final trend reported is related to the "I forget" response. 15% or more of the preservice teachers responded "I forget" to 17 of the 83 items of thePTPQ. Of the 17 items, at least 9 items reflected abstract concepts in regard to reading instruction. The abstract concepts primarily consisted of defining attributes of reading and rationales for a particular reading activity.

In general, the less frequently perceived teacher behaviors – respective to trends four, five, and six, and abstract behaviors respective to trend seven, are related to the reasons for, or the why of, reading instruction.

Conclusions

The results, respective to both the characteristics of the instrument and the information resulting from its application, have been generally encouraging. If the assumption holds that the teaching behaviors perceived to occur more frequently exert greater influence than those teacher behaviors that are of lesser frequencyl, then the results obtained seem in accordance with the differences observed between preservice teacher performance and suggested performance derived from information within a course of study. Specifically, if the stronger perceptions of preservice teachers, respective to practical application of instruction, exert a greater influence than the weaker perceptions-purposeness and rationale for instruction, then the observed differences between actual and suggested performance respective to inappropriate instruction seem interpretable. The resulting interpretation would then suggest that preservice teachers have clear perceptions as to how to provide instruction but unclear perception as to why certain instructional practices are in order. Further research is needed, of course, in order to test this interpretation.

References

Balzer, A. L., Evans, T. P., & Blosser, P. E. *A review of research on teacher behavior.* Columbus, Ohio: ERIC information Analysis Center for Science, Mathematics, and Environmental Education, December 1973.

Rosenshine, B. & Furst, N. Research on teacher performance criteria. In B. O. Smith (Ed.), *Research in teacher education: A symposium.* Englewood Cliffs, New Jersey: Prentice-Hall, 1971.

Smith, G. *Request for a pilot operating grant-Triple program.* Unpublished proposal, Indiana University, January 1969.

EUNICE N. ASKOV
MARY M. DUPUIS
Pennsylvania State University

JOYCE W. LEE
State College Area School District

An effective inservice model for content area reading in the secondary schools

Staff development programs for teachers may take many forms as have been described in recent literature. Often typical inservice programs, described as "various kinds of one-shot 'dog and pony shows' " (Dillon, 1976, p. 169), involve a visiting consultant who is invited to the school for an hour or two to speak on his/her area of expertise. Teachers listen dutifully; some even take notes. But no serious change in their practices and attitudes is either expected or achieved. The problem with this "haphazard" approach to staff development is that "it is disorganized, piecemeal, patchwork" (Edelfelt, 1977, p. 112). This paper is a report of the development of a model for reading inservice education in the secondary schools that was designed to have an impact.

The Content Area Reading Project, which was funded by the Pennsylvania Department of Education from January 1, 1976 - June 30, 1977, involved an inservice effort for junior high school and adult basic education/high school equivalency (ABE/GED) teachers. The Project's goal was, first, to change attitudes positively toward teaching reading as part of content area subjects and, second, to effect change in that direction in classroom practices. Although all cooperating schools had identified content area reading as a weakness among students primarily due to inadequate teaching strategies, we realized that as outsiders or "visiting experts" from a university, we might, in fact, have very little impact on teachers' attitudes and classroom practices (Otto & Erickson, 1973). Therefore, we chose a model for inservice education which might be used more often by a consultant regularly employed by a school district (e.g., curriculum director, reading consultant) in which a series of inservice sessions could be scheduled over a long period of time with follow-up in the classrooms.

The school administrators, particularly at the junior high school level, were well aware of the problems of integrating reading with content subjects. Content teachers who may be well versed in their subject areas frequently have difficulty individualizing instruction for different reading abilities due to large, changing content classes (Burnett & Schnell, 1975). Olivero (1976) also suggests that "probably less than one percent of the secondary teachers have ever been taught the skills" (p. 195) of diagnosis necessary for individualization of instruction. They may simply not have had any course work in reading methods during their own preservice teacher training experience (Morrison & Austin, 1977) and may not be aware of how to help students read content materials. Or they may resist teaching reading skills "by rationalizing that 'reading isn't my subject' " (Axelrod, 1975, p. 82).

Therefore, as Olivero (1976) suggests, changing attitudes was considered of primary importance. Since affective objectives may be attained primarily through activities with high experience impact and two-way communication (Otto & Erickson, 1973), follow-up between inservice sessions by graduate assistants serving as reading consultants seemed essential to observe and provide feedback on progress toward integrating reading skills in the content subject.

A competency-based format was selected since differences among the teachers were anticipated in entry levels and progress in both attitudes and skills. A computer record-keeping system was also employed to keep track of teachers' progress in meeting workshop objectives or competencies. At each session teachers received a computer printout showing objectives mastered and unmastered. We hoped that modeling a diagnostic-prescriptive approach in the workshops would influence teachers toward diagnostic-prescriptive teaching. Objectives, which are presented elsewhere (Dupuis & Askov, 1977), were written at three levels: cognitive, simulation, and application. It was anticipated that the skills learned in the workshops would be applied and observed in the classroom where the teacher integrates the skills previously mastered at the cognitive and simulation levels.

Method

Subjects

The target group for the study consisted of teachers of junior high and adult students at three sites. Since reading instruction for ABE/GED teachers has been generally lacking (Hall & Coley, 1975), they were included with junior high school teachers who teach potential or future members of ABE/GED classes (for students who drop out of school). Reading specialists were also included so that they might become more sensitive to the demands of the content area classroom and provide help to content teachers as they deal with students of all reading abilities.

Procedure

Following recommendations made by others (Campbell, 1973, Parker & Campbell, 1973), we planned a year-long program, consisting of fifteen bimonthly sessions, each lasting three hours, for volunteer teachers. Three or six hours of university credit were available as an option for teachers who wished to register for the credits.

Three sites were selected to represent urban, suburban, and rural settings. The workshops were held at the junior high school in each site. Teachers were offered the option of released time during the school day with substitutes hired by the project; however, they instead elected to receive extra compensation for attending the training sessions in the evening or after school.

A staff of four graduate education students was hired to serve as school consultants. Each was assigned to teachers in a given school to provide follow-up in the classrooms between workshop sessions. They not only attended the workshops at their particular sites, assisting the two university instructors, but also observed classes, held conferences with teachers, and did demonstration teaching as requested. They served in the role of "helper" (Long, 1977, p. 81) and facilitator of change.

Materials

Model materials were created by project staff to demonstrate how teachers might construct and use these materials in their classrooms. We particularly hoped by modeling the use of multimedia materials in the workshops to encourage the use of these materials in the classrooms. For example, a videotape of a vocabulary/social studies lesson was created which demonstrated how a content area teacher could make his/her own videotape to be used in classroom instruction. A listening center pertaining to note-taking was set up to give teachers experience in listening comprehension and to encourage the use of such centers in their classrooms.

An advisory board, consisting of professors at The Pennsylvania State University from the various content areas which are typically taught in junior high schools and adult programs, was established. Each advisory board member identified sample materials from his/her content field that might be used in junior high school and ABE/GED programs as well as professionsl books relating reading skills to the content area subject. A professional library, consisting of books suggested by advisory board members plus many other resources dealing with content area reading, was developed for each site. The professional library was provided so that teachers could complete assigned readings without having to travel to a university library.

Evaluation

An evaluation design was planned to determine the effectiveness of the inservice education mode. Following suggestions for evaluating staff development programs (Bishop, 1977), various types of formative and summative evaluation were planned in addition to assessing mastery of each workshop objective. Since results could not be expected among junior high and ABE/GED students immediately, only teacher variables were considered. So few teachers of adults signed up for the project that their data were excluded from the following analyses. Therefore, conclusions are presented only in terms of the junior high school teachers.

Since attitude change was considered to be a prerequisite to changes in classroom practices, attitudes toward incorporating reading skills in content area subjects were assessed with two instruments, a twenty-item Likert scale (Statements Survey, $r = .84$) and a less direct instrument in which teachers responded to descriptions of twelve teaching situations using five sets of bipolar objectives on a semantic differential scale (Situations Survey, $r = .90$). (See Dupuis & Askov, 1977, for information on the development of the instruments.) Workshop teachers were pre and posttested with these two attitude instruments and the Purdue Teacher Opinionaire (Bentley & Rempel, 1973). The Purdue Teacher Opinionaire (PTO), a measure of satisfaction with the teaching situation, was administered to control for changes in attitudes that might be due to school factors rather than to the Project. A comparison group, teachers who worked in the same schools but were not part of the workshops, was established in each junior high school; comparison group teachers were administered the same pre and posttests.

A treatment x time interaction effect was found, $F(1,123) = 5.46$, $p < .05$, with experimental teachers' gains on the attitude measures significantly greater than the gains of the comparison teachers. No significant time effect was found for the morale measure (PTO); morale, although significantly lower at the urban site, remained constant at all sites and did not appear to be a significant factor in determining attitude toward incorporating reading instruction in content area classrooms. The means are presented in Table 1.

In addition to assessing changes in teachers' attitudes, a criterion-referenced knowledge of reading skills test, which measured teachers' mastery of the cognitive aspects of the workshop objectives, was administered only to workshop teachers at the beginning of the first workshop session and again at the last workshop session. Mastery level was established at 80%. Of the 56 experimental group teachers who took the test at the first workshop session, only three scored 80% or better. Of these same 56 teachers, significantly

Table 1

Pre and Posttest Observed Mean Scores on Attitude Measures

	All Sites Combined		Urban Site Only		Suburban Site Only		Rural Site Only	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Treatment Groups Combined:								
Statements Survey Scores (Likert scale)	85.40 (*n* = 129)	87.28 (*n* = 129)	83.78 (*n* = 41)	86.29 (*n* = 41)	88.29 (*n* = 49)	88.59 (*n* = 49)	83.49 (*n* = 39)	86.67 (*n* = 39)
Situations Survey Scores (Semantic differential scale)	348.17 (*n* = 129)	366.68 (*n* = 130)	349.76 (*n* = 42)	358.40 (*n* = 42)	354.69 (*n* = 48)	373.69 (*n* = 49)	338.44 (*n* = 39)	366.77 (*n* = 39)
PTO	288.90 (*n* = 127)	288.11 (*n* = 127)	235.08 (*n* = 40)	231.95 (*n* = 40)	323.57 (*n* = 49)	324.06 (*n* = 49)	300.84 (*n* = 38)	300.86 (*n* = 38)
Experimental Groups Only:								
Statements Survey Scores	86.06 (*n* = 58)	90.78 (*n* = 58)	84.62 (*n* = 21)	91.81 (*n* = 21)	88.90 (*n* = 20)	92.05 (*n* = 20)	84.53 (*n* = 17)	88.00 (*n* = 17)
Situations Survey Scores	346.96 (*n* = 57)	374.90 (*n* = 57)	347.43 (*n* = 21)	368.95 (*n* = 21)	351.68 (*n* = 19)	377.00 (*n* = 19)	341.12 (*n* = 17)	379.38 (*n* = 17)
PTO	284.36 (*n* = 55)	284.33 (*n* = 55)	245.21 (*n* = 19)	239.68 (*n* = 19)	315.00 (*n* = 20)	317.70 (*n* = 20)	292.63 (*n* = 16)	295.63 (*n* = 16)
Comparison Groups Only:								
Statements Survey Scores	84.86 (*n* = 71)	84.10 (*n* = 72)	82.90 (*n* = 20)	79.57 (*n* = 21)	87.86 (*n* = 29)	86.21 (*n* = 29)	82.68 (*n* = 22)	85.64 (*n* = 22)
Situations Survey Scores	349.13 (*n* = 72)	360.25 (*n* = 72)	352.10 (*n* = 21)	347.86 (*n* = 21)	356.66 (*n* = 29)	371.97 (*n* = 29)	336.36 (*n* = 22)	356.64 (*n* = 22)
PTO	293.50 (*n* = 70)	292.93 (*n* = 70)	223.15 (*n* = 19)	225.11 (*n* = 19)	329.48 (*n* = 29)	328.45 (*n* = 29)	306.82 (*n* = 22)	304.68 (*n* = 22)

more teachers, 22, reached mastery level of 80% on the posttest given at the final workshop session. In addition, in a site x time analysis of variance of actual pre and posttest scores, significant gains, F(1,53 = 196.02, p .001, were found. This would indicate that workshop participants significantly increased their knowledge of basic methods for teaching reading in content areas.

The graduate assistants assigned to work onsite with the teachers also provided ratings of the experimental teachers at the beginning and end of the school year. These ratings pertained to the degree to which the teachers were judged to have incorporated reading skills in their classrooms. Mean exit ratings were significantly higher, $t(57) = 9.851$, $p < .001$, across all sites (Dupuis & Askov, 1977). Therefore, it appears that what was being learned in the workshops was indeed being applied in the classrooms.

The data, as summarized, seem to indicate that the year-long inservice program in content area reading significantly changed junior high school teachers' attitudes toward and knowledge of reading skills. Furthermore, the consultant ratings indicate that these changes were reflected in classroom practices; teachers were observed to incorporate reading instruction into the content study to a greater extent after the treatment period. Classroom application was considered an important aspect of the Content Area Reading Project since most inservice efforts stop short of classroom observations to make sure that workshop learnings are being applied during instruction.

In addition to the summative evaluation data just reported, other kinds of data were gathered in project evaluation. The issue of whether university credit had an impact on mastery of workshop objectives was considered since credit is often considered as a necessary incentive in inservice teacher education. Of the total experimental sample, 60% elected to register for graduate credit while participating in the workshop program: 86% of the teachers at the urban site, 30% at the suburban site, and 65% at the rural site. Of the experimental teachers registered for credit, 89% satisfactorily completed all objectives required for graduate credit: 83% at the urban site, 100% at the suburban and rural sites. Of those who did not register for credit, 13% satisfactorily completed enough objectives to earn a certificate of completion (awarded to those who completed the required workshop objectives).

However, when credit status (or credit x time) was examined in relation to attitude scores, no significant effects were found. Whether or not a teacher elected to participate in the project for credit or no credit was not significantly related to attitude change during the experimental treatment period.

Furthermore, no significant effects were found for the credit status (or credit x time) on the criterion-referenced knowledge of reading skills test. In other words, change in the criterion-referenced test scores was not significantly related to whether or not a teacher elected the graduate credit program option offered as part of the experimental treatment.

Therefore, although credit status did affect the number of workshop objectives completed, it did not seem to be related to changes in attitude toward and knowledge of reading skills. Apparently, changes in attitudes and knowledge were occurring from workshop attendance alone rather than from completing the required work.

Evaluation data provided by the teachers in terms of the value and usability of each workshop activity are reported elsewhere (Dupuis & Askov, 1977). In general, teachers tended to favor the practical activities with direct classroom application (e.g., vocabulary exercises) as opposed to those that were more theoretical in nature (e.g., the concept of linguistic differences).

They were also extremely positive toward the follow-up help given between workshops by the graduate assistants.

Teachers also responded generally favorably to the competency-based format of the inservice program; however, a period of adjustment to the idea of competencies should have been built into the time frame. Teachers were inexperienced with mastery learning, especially with the concept of multiple revisions until satisfactory work is attained. Nevertheless, Project staff felt that CBTE was effective in working with inservice teachers. If nothing else, it modeled the process of mastery learning which teachers now can apply in their own classrooms. Computer record-keeping was considered valuable by both staff and teachers in keeping track of work submitted and work needing revision.

One of the most perplexing problems was the poor response by ABE/GED teachers in joining the workshops. In a questionaire sent to those who had originally expressed interest but did not sign up for the workshops, we discovered that time appeared to be the biggest problem. Since the workshops were scheduled after the traditional school day, they conflicted with teaching schedules of teachers of adults whose classes were held primarily in the afternoons and evenings. Many of these teachers already held two jobs, not leaving much time for inservice work. It was suggested that a specially designed inservice program (perhaps through packaged materials and videotapes) should be delivered at the site of the adult learning center for small groups of teachers to provide flexibility in scheduling inservice work. This type of inservice model specifically for teachers of adults has been discussed elsewhere (Schroeder & Haggerty, 1976).

Conclusion

The inservice model appeared to be effective in bringing about changes in attitudes toward and knowledge of reading skills among junior high school teachers. Evidence also exists that these changes were having an impact on the classroom instruction of the experimental teachers. The critical ingredients appeared to be that the workshops were delivered at the junior high schools instead of in university classrooms and that follow-up assistance was available onsite between workshop sessions to help teachers apply what they had learned in their own classrooms. Perhaps impact would not have been so great if the teachers had not been forced by the structure of the required objectives to apply theory and techniques learned in workshops in their classrooms. Whether or not these changes affect student achievement, which, of course, is the ultimate goal, can be determined only by follow-up research. Positive changes in attitude and increased knowledge of reading skills, however, are important prerequisites to better reading instruction in the content areas.

References

Axelrod, J. A few recommendations on how to conduct inservice reading instruction for content area teachers. *English Journal*, 1975, *64*, 81-82.

Bentley, R. R. & Rempel, A. M. *Purdue teacher opinionaire*. Lafayette, Indiana: Purdue Research Foundation, 1973.

Bishop, L. Visualizing a staff development plan. In C. W. Beegle & R. A. Edelfelt (Eds.), *Staff development: Staff liberation*. Washington, D.C.: Association for Supervision and Curriculum Development, 1977.

Burnett, R. W. & Schnell, T. R. CBTE and secondary reading: A philosophy and rationale. *Journal of Reading*, 1975, *18*, 544-549.

Campbell, J. J. A field-centered, competency-based inservice reading program at an inner city school. In P. L. Nacke (Ed.), *Interaction: Research and practices in college-adult reading.* Milwaukee: National Reading Conference, 1973.

Dillon, E. A. Staff development: Bright hope or empty promise? *Educational Leadership*, 1976, *34*, 165-170.

Dupuis, M. M. & Askov, E. N. *The content area reading project: An inservice education program for junior high school teachers and teachers of adults* (Final Report, Project 09-6905). University Park, PA: The Pennsylvania State University, 1977.

Edelfelt, R. A. Staff development: What's to be done in the future? In C. W. Beegle & R. A. Edelfelt (Eds.), *Staff development: Staff liberation.* Washington, D.C.: Association for Supervision and Curriculum Development, 1977.

Hall, M. & Coley, J. D. Needs in reading instruction in adult basic education. *Adult Leadership*, 1975, *24*, 103-104.

Long, D. H. Competency-based staff development. In C. W. Beegle & R. A. Edelfeldt (Eds.), *Staff development: Staff liberation.* Washington D.C.: Association for Supervision and Curriculum Development, 1977.

Morrison, C. & Austin, M. C. *The torch lighters revisited.* Newark, Delaware: International Reading Association, 1977.

Olivero, J. L. Helping teachers grow professionally. *Educational Leadership*, 1976, *34*, 194-200.

Otto, W. & Erickson, L. *Inservice education to improve reading instruction.* Newark, Delaware: International Reading Association, 1973.

Parker, B. & Campbell, A. B. A new approach to graduate education in adult basic and GED education. *Adult Leadership*, 1973, *22*, 208-209.

Schroeder, S. A. & Haggerty, D. M. Competency-based individualized training programs for adult educators: A non-traditional approach to adult staff development. In G. E. Spear (Ed.), *Adult education staff development: Selected issues, alternatives and implications.* Kansas City, MO: Center for Resource Development in Adult Education, 1976.

CAROL A. DAVIS
Purdue University

The effectiveness of informal assessment questions constructed by secondary teachers

The documentation and research concerning informal reading inventories has dealt almost exclusively with elementary school use. Few studies have been done on the use or effectiveness of these measures at the secondary level. Despite this lack of documentation, numerous reading authorities have advocated the use of informal reading inventories at the secondary level (Froese, 1974). Teachers have been frequently encouraged to develop their own informal assessment techniques. Teacher-constructed inventories have been favored because a minimal amount of work has been done on published inventories with secondary level passages or on inventories which meet the specialized needs of content areas. In an effort to aid teachers with this task, articles have been written describing effective questioning procedures and constructions (Valmont, 1972; Lucking, 1975). Studies have been undertaken to improve the training of teachers in questioning techniques (Rogers and Davis, 1970; Trosky, 1971). In spite of these efforts, the vast majority of secondary teachers have not been adequately prepared for developing informal assessment questions. The appropriateness or effectiveness of informal reading inventories constructed by these untrained teachers has not been determined. It is the purpose of this paper to suggest possible procedures for evaluating the effectiveness of teacher-constructed informal reading inventory questions and to apply these procedures to an experimentally generated set of questions. The encouragement of secondary teachers to develop their own IRIs will be critically examined.

Method

Materials

The informal reading inventory used in the study was developed to provide a reasonably quick, yet effective means of assessing the reading performance of secondary students individually or as a group (Childrey, 1977). The inventory passages were selected by experienced teachers at the secondary level. The passages included material from four areas: English/reading, social studies, science, and a miscellaneous group sampling such areas as vocational arts, fine arts or mathematics. The passages ranged from 260 to 350 words in length. Each of the passages was evaluated by means of the Fry readability graph and the SMOG formula. Two passages each at readability grade levels five, seven, nine, eleven, and thirteen were chosen for use in the study. An effort was made to select passages on the basis of potential as high interest material. Each passage was accompanied by a set of teacher-constructed questions consisting of two vocabulary, three knowledge, three inference, one speculation, and one application question.

Procedure

The passages and questions were administered to seventh, ninth, and eleventh grade students in an upper middle class suburban community located near a large city. The inventory was administered as a group task by the classroom teachers. All inventories were scored by a single individual.

Design

A three factor design with two between-subjects variables, grade and passage, and one within-subjects variable, question types, was used in the study. Each student read passages at one grade level.

Results

The number of correct responses for each set of the teacher-constructed questions was tabulated and then subjected to a three-way analysis of variance. Significant effects were found for passage, F (4, 165) = 5.81, $p <$ .01, and question type, F (3, 495) = 20.18, $p < .001$, and for interactions between grade and passages, F (8, 165) = 2.45, $p < .05$, and passages and question types, F (12, 495) = 3.87, $p < .01$. A significant effect was not found for the grade variable. However, linear trend analyses for grade and passage indicated the expected progression across grade levels and across passages. Tests carried out using the Newman-Keuls procedure on the question variable indicated vocabulary questions were significantly more difficult than other question types.

The questions were next evaluated by use of a discrimination and difficulty index program (Linden & Mazzucca, 1977). The difficulty index was identified as the percentage of subjects who answered the questions correctly. Questions were subsequently labeled as easy, average, or hard. The discrimination index was identified as the ability of the questions to discriminate between high and low scoring subjects and represented the percentage of maximum possible discriminations obtained. Questions were identified with ratings varying in range from negatively discriminating to highly discriminating.

The results of the difficulty and discrimination index program were used to determine whether or not the questions had been effective with subjects at the grade level they were intended for. If exhibiting appropriate difficulty at grade level, questions were expected to consist of a large portion of average items with a smaller balance of easy and hard items. The results of this analysis indicated an acceptable balance of difficulty for the inventory questions. For seventh grade level the results showed three easy, six average, and one hard question. The ninth and eleventh grade results showed one easy, eight average, and one hard question. If discriminating effectively, questions were expected to lie within the plus one to plus four discriminating range, preferably at plus two or above. The program results indicated that while there were no negatively discriminating items, the bulk of questions were non-discriminating or low discriminating items with a plus one rating.

The results of the program were also used to determine whether or not logical progression in terms of difficulty and discrimination existed across grade levels. It was expected, for example, that questions accompanying fifth grade level passages would be easy and nondiscriminating for the seventh, ninth, and eleventh grade subjects. The program results indicated the questions were of appropriate difficulty across grade levels but were inadequate or inconsistent in progression and levels of discrimination.

Discussion

The analysis of variance results suggest that as a whole, the set of inventory questions operates appropriately by demonstrating expected differences among the subjects and the graded passages and questions. However, the other evaluative procedures suggest the questions are not effectively constructed. The Newman-Keuls tests indicate teachers have difficulty developing appropriate vocabulary questions. Although the difficulty level of questions appears satisfactory both for intended grade level and across grade levels, the inventory questions often do not discriminate between high and low scoring subjects. Certainly only questionable use can be made of an informal reading inventory if its items cannot discriminate between good and poor readers.

The results of this study seem to suggest the need for reconsideration in encouraging teacher-constructed secondary reading inventories. Pre-service and inservice training programs need to focus on better preparation for teachers in constructing and evaluating questions. More thorough instruction in the use, scoring, and interpretation of informal reading inventories is needed. Teachers must be repeatedly cautioned to use such results in conjunction with other available pupil information. Encouragement to create informal reading inventories should only be given when adequate preparation for such a task has taken place.

If used appropriately by a trained professional, teacher-constructed assessment questions can work effectively. However, it seems possible that inappropriate assumptions have been made in encouraging secondary teachers to construct IRIs.

References

Childrey, J. A. Informal reading inventories for teachers of secondary school subjects. *Remedial Education*, 1977, *12*, 74-78.

Froese, V. IRIs at the secondary level re-examined. In P. Nacke (Ed.), *Twenty-third Yearbook of the National Reading Converence*. Clemson, S. C.: National Reading Conference, 1974.

Linden, K. W. & Mazzucca, S. A. *ITANAL: Item analysis and weighted item scoring computer program*. West Lafayette, Indiana: Purdue Research Foundation, 1977.

Lucking, R. A. *Comprehension and a model for questioning*. Paper presented at the Annual Secondary School English Conference of the NCTE, Kansas City, April 1975.

Rogers, V. & Davis, O. L. *Varying the cognitive levels of classroom questions: An analysis of student teachers' questions and pupil achievement in elementary social studies*. Paper presented to the American Educational Research Association, Minnesota, March 1970.

Trosky, O. A. *Modifications in teachers' questioning behavior in the development of reading comprehension and a series of supervisory conferences*. Paper presented to the International Reading Association, Atlantic City, April 1971.

Valmont, W. Creating questions for informal reading inventories. *Reading Teacher*, 1972, *25*, 509.

JEFFREY L. LORENTZ
HOMER COKER
West Georgia College

Observed patterns of teacher-pupil classroom behavior as predictors of student growth in reading

This paper explores the relationships between observed classroom behavior and student achievement in reading. In two earlier studies, the investigators reported significant relationships between classroom mean student coping style, teacher control, and reading achievement (Coker and Lorentz, 1975), as well as between individual student coping style and reading achievement (Coker and Lorentz, 1976). The present paper is a further examination of data collected in 1974-76 as part of the Carroll County CBTC Project. It is the purpose of the present study to examine the relationship between observed classroom behavior, specifically, teacher-pupil interactions and reading achievement.

It was hypothesized that certain teacher-pupil interactions, as recorded by direct observation using the *Spaulding Teacher Activity Rating Schedule* would be significant predictors of end of year reading achievement.

Method

Subjects

The participants in the present study, 41 classroom teachers, grades 3-8, are a subset of the population of 103 teachers who participated in the CBTC Project during the 1974-76 school year (1974-75 sample = 26 classes, 1975-76 sample = 15 additional classes). Only those teachers who taught reading as a major subject are included in the sample selected for the present study.

An estimate of the socioeconomic status (SES) was obtained for each pupil based upon the occupation of the family's principal bread-winner. This occupational information was transformed to a 1-8 scale (1 = high status, 8 = low status) using a modified version of the Warner, Meeker, and Eells (1960) classification of occupations and levels. Pupils with missing SES were assigned the mean for their classroom, rounded to the nearest whole number.

Materials

The *Iowa Tests of Basic Skills*, Form 5 and 6, Levels Edition, (ITBS) (Hieronymus and Lindquist, 1971) was given to all pupils in grades 3 through 8 who were involved in the CBTC Project. The *Reading Comprehension* subtest only was used in the present study. The reported reliability and validity for the ITBS are adequate for use in the present study. In addition, the State of Georgia used the ITBS in the 4th and 8th grades in its state-wide testing program.

The *Spaulding Teacher Activity Rating Schedule* (STARS) (Spaulding, 1975) is a category system which examines the cognitive instructional strategies of teachers as well as their affective and control techniques. STARS consists of 25 categories of teacher behavior which are subsumed under the subtitles of affective behavior, motor and social structuring, concept attainment, concept checking, and value expression, as they interact with 19 categories of student behavior which are identified by descriptive statements such as "Aggressive

Behavior", "Self-directed Activity", "Observing Passively", etc. STARS has been developed over a period of ten years and is currently being used in a number of studies.

Procedure

The initial reading tests were administered early in the fall in the classrooms of the 41 Carroll County teachers, and posttests were administered in late spring in the same classrooms. A group of specially-trained testers administered the tests in each classroom.

After the pretests had been administered, each of the approximately 1200 pupils in the 41 classes was observed on two separate visits during a three-week period by observers using the CASES observation instrument (Spaulding, 1970). Each pupil's record was scored to identify a predominant coping style, such as Aggressive, Passive, Task-oriented, etc.

After the pretests had been administered and the CASES scored, six pupils with different predominant coping styles were randomly selected from each classroom. These six students were used in subsequent observations using STARS. Six such observations were made in each classroom during the remainder of the school year.

Visits were scheduled in advance, but no attempt was made to preselect the activities to be observed, since the intent was to record a representative sample of the behavior present in each classroom.

All observations were made by experienced classroom teachers who had been employed full-time by the CBTC Project as observers and had been personally trained by the author of CASES and STARS, Dr. Robert Spaulding. Observer agreement studies, conducted in non-project classrooms, confirmed that the observers were consistently able to maintain an agreement of 80 percent or better.

Analysis

Tests were hand or machine scored, keypunched and verified. Observation data were keypunched from the data collection forms and verified.

ITBS scores were converted from raw scores to Standard Scores using the published Standard Score Tables (Houghton Mifflin, 1973). These scores used a normalized standard score scale with a mean of 80 and a standard deviation of 20 for the entire grade range, 3-8 data in a single analysis.

STARS frequencies for each pupil were combined using Spaulding's procedure (Spaulding, personal communication) to identify twenty-one scores. These are derived from combinations of behaviors recoreded in the 475 cell (i.e., 25 x 19) student/teacher interaction matrix and labeled with titles descriptive of individual teaching behaviors.

The pretest and posttest reading scores, a SES measure, grade level, and twenty-one STARS scores were the variables used.

Two additional variables were generated for each student. These were Pre x Grade and Pretest Squared. Pre x Grade was generated because the correlations of pretest and posttest tend to increase at higher grade levels, and this term permits fitting different regression slopes at different grade levels. The Pretest Squared term permits fitting a two degree curve to the relation between pretest and posttest, and tests if a nonlinear fit is significantly better than a linear one.

To examine the relationships between the STARS scores and reading achievement, two stepwise multiple regression analyses were carried out using the SPSS (Nie et al., 1975) regression procedure.

The dependent variable in both analyses was posttest. In the first analysis, the independent variables were forced to enter the equation in the following order: pretest, grade, SES, Pre x Grade, pretest squared, and finally STARS 1 through STARS 21. In the second analysis, pretest was forced to enter first and the remaining variables were permitted to enter in the order in which they accounted for additional proportions of the variance.

Results

As predicted, pretest is the best predictor of posttest Reading (r = .81), accounting for nearly two-thirds of the variance. Several STARS scores were better predictors of posttest reading than were grade and SES. Therefore, the second analysis, in which all variables were permitted to enter the equation depending on their contribution, is considered in the following discussion. Table 1 presents the variables in their order of entry into the equation.

Table 1

Multiple Regression Analysis

	Variable	Multiple R	Simple R	F to Enter	
	Pre-Reading	81	81	376.18	***
21.	Structuring Supervisor	82	-23	10.01	**
	Pre x Grade	83	78	8.46	**
19.	Supportive Guide	83	-20	6.48	*
	SES	84	-27	5.41	*
20.	Egocentric Instructor	84	32	4.65	*
2.	Boring Lecturer	84	15	2.77	
17.	Style EFG Treatment	85	-03	2.32	
10.	Effective Manager	85	-01	5.09	*
13.	Style A Treatment	85	02	1.42	
	Pretest Squared	85	81	1.07	
	Grade	85	64	3.21	
3.	Examiner	86	-18	1.64	
6.	Counselor	86	-07	1.48	
4.	Entertainer	86	32	0.70	
12.	Rote Process	86	-16	0.33	
8.	Discovery	86	-12	0.46	
9.	Socratic	86	06	0.62	
15.	Style C Treatment	86	-09	0.20	
1.	Effective Story Teller	86	19	0.19	
11.	Expository	86	08	0.27	
5.	Controller	86	-07	0.09	
7.	Pseudo Peer	86	15	0.09	
14.	Style B Treatment	86	-03	0.02	
18.	Responsive Explainer	86	07	0.01	

*** $p < .001$
** $p < .01$
* $p < .05$

Discussion

One unexpected result was that several of the STARS scores were better predictors of reading achievement than were either Grade Level or SES. Another surprising finding was the negative relationship between reading and teacher/pupil interactions judged to be "good", such as Structuring Supervisor and Supportive Guide. Also unpredicted was the positive relationship between Reading Achievement and Egocentric Instructor.

The lack of significant contribution of other "good" STARS scores such as Discovery Teacher and Socratic Teacher is probably due to the low frequency of these behaviors in the Carroll County data.

These findings are in general agreement with other studies. They suggest that while the best indicator of a student's growth in reading is his pretest score, observed behaviors of both students and teachers may be useful predictors of achievement in reading.

References

Coker, H., & Lorentz, J. L. *An examination of student coping style, teacher control, and student achievement in reading.* Paper presented at the 25th Annual Meeting, National Reading Conference, St. Petersburg, Florida, 1975.

Coker, H., & Lorentz, J. L. *Growth in reading as a correlate of pupil classroom behavior.* Paper presented at the 26th Annual Meeting, National Reading Conference, Atlanta, Georgia, 1976.

Hieronymus, A. N., & Lindquist, E. F. *Iowa tests of Basic skills.* Boston: Houghton Mifflin, 1971.

Iowa tests of Basic Skills, Forms 5 and 6. *Age equivalent/standard score tables.* Boston: Houghton Mifflin, 1973.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. *Statistical package for the social sciences* (2nd ed.). New York: McGraw-Hill, 1975.

Spaulding, R. L. *Classroom behavior analysis and treatment.* San Jose: Author, 1975.

Spaulding, R. L. *Spaulding teacher activity rating schedule.* San Jose: Author, 1975.

Warner, W. L., Meeker, M., & Eells, K. *Social class in America.* New York: Harper, 1960.

CHRISTOPHER J. RAMIG
ROBERT C. GRANGER
Georgia State University

Teacher inferences of the characteristics of non-standard speaking readers

Early attempts to teach reading to non-standard Black Dialect speakers were based for the most part on the assumption that the language used by these children was deficient or incomplete in development. This was due, it was proposed, to the culturally deprived environment from which these children came; a lack of early stimulation resulted in inadequate speech and retarded cognitive development and subsequent failure in school, especially in learning to read.

Challenges to this deficit model of language development began with McDavid's (1964) plea for teacher acceptance of a concept of Black Dialect as a *different* language system rather than a *deficient*, underdeveloped version of Standard English. According to most sociolinguists (e.g., Baratz, 1969), Black Dialects are well-formed, highly developed linguistic systems with phonological and grammatical rules of their own. While different from Standard English, the dialects are not deformed or deficient versions of Standard English.

The deficit model "easily" explained reading failure as lack of language and cognitive development, and recommended remediation in language, i.e., standardization from "amalgamated noises" to Standard English phonology and grammare (Bereiter and Englemann, 1966). The difference model explained reading failure as a "mismatch" between the linguistic system of the Black Dialect speaking child and the linguistic system represented in beginning reading materials (Goodman, 1969).

The mismatch hypothesis led to a number of research projects conducted to assess the effects of dialect difference on reading comprehension and on learning to read (Nolen, 1972; Sims, 1972; Baratz, 1973; Melmed, 1973; Hockman, 1974). Representative of this research is a study by Simons and Johnson (1974) in which it was postulated that Black Dialect speaking children would comprehend passages written in Black Dialect better than they would comprehend passages written in Standard English. Black Dialect versions of Standard English materials were written and presented to Black Dialect speaking children. No evidence was found to support this version of the dialect interference theory.

Findings such as this have led to other "explanations" of the reading difficulty many children experience. One of these is the argument that a potential cause of the reading failure of nonstandard speaking children is a tendency of teachers to force word-for-word accuracy in oral reading, rather than viewing reading as a meaning getting process (Goodman and Buck, 1973; Cunningham, 1975). The latter conception of reading would require a teacher to distinguish between meaning versus non-meaning changing miscues or deviations. The distinction is important since Goodman and Sims (1974) report that "Black Dialect speakers frequently read Standard English structures orally as Black Dialect structures" and that "changes made by the subjects were surface changes (and) retained the meaning of the original sentences" (p. 838).

Previous research findings suggest that teachers consider the speakers of nonstandard dialects to be less adequate in their speech than speakers of

Standard English (Naremore, 1971; Whitehead and Miller, 1972; Williams, Whitehead and Miller, 1971; Granger, Quay, Mathews and Verner, 1977). There are also several studies suggesting that there is a relationship between the dialect of a speaker and responses toward that speaker. Gess (1969) and Steadman and Adams (1973) have found a high positive correlation between speech and ratings of student behavior. Holmes (1968), Guskin (1970), Crowl and MacGinitie (1966) and Covington (1972) have all found positive correlations between ratings of speech and predictions of academic test scores, intelligence, and academic readiness.

There is from the research, then, a suggestion that teachers' response to nonstandard dialect may affect their judgments of children's behavior and academic ability. Such judgments may affect instructional procedures such as those often found in reading instruction, manifesting in what Goodman and Buck (1973) have termed linguistic discrimination.

The empirical evidence for such a conclusion is, however, sketchy. It was the purpose of the present study to determine whether or not the presence of Black Dialect syntactic features in the speech of readers in oral reading situations would affect teacher perception of the reader's reading ability. Do teachers consider readers who produce dialect based non-meaning changing miscues as less able readers than their Standard English speaking counterparts? If so, is this equally true for good and poor readers?

Method

Subjects

The subjects were 84 students enrolled in three different sections of a graduate course on methods and materials for reading in the elementary school. The sample contained 68 whites, 15 blacks and one Asian student. There were four males and 80 females. Seventy-four of the subjects were teaching or had taught, with a mean teaching experience of 4.7 years. Of the teachers, 48 had been or were currently, directly responsible for teaching reading.

Materials

Using a story by Clark (1966) the investigators constructed four variations of the text, each reflecting a different type of reader: Standard English good reader (SEGR), Black Dialect good reader (BDGR), Standard English poor reader (SEPR), and Black Dialect poor reader (BDPR). Each variation of the text contained fifteen changes from the original but following the distinction between meaning-changing miscues (MCM) and non-meaning-changing miscues (NMCM), good reader versions (SEGR and BDPR) and poor reader versions were prepared. SEGR and BDGR versions contained five MCM and ten NMCM while the poor reader versions (SEPR and BDPR) contained ten MCM and five NMCM. Of the NMCM, all versions shared five, with the second five for SEGR typical of a Standard English speaking good reader's miscues and the second five for BDGR indicative of Black Dialect syntactic features (e.g., double negation, progressive *be*).

Three "readings' of each version were then audio recorded by adult females–three Standard English speakers and three Black Dialect speakers. Each of these six adult females read both a good reader version and a poor reader version of the text. Three sets of four readings were then created with each reader type present in each set but with order within a set randomized and no good reader and poor reader version as recorded by the same adult female present in the same set. Thus, no subject listened to a good reader

31

Table 1

Miscues Across Versions

		Reader	
		Good	Poor
Dialect	Standard English	MCM = 5[1] NMCM = 5[2] NMCM = 5[2] Total Miscues = 15[5]	MCM = 5[1] NMCM = 5[2] MCM = 5[4] Total Miscues = 15[5]
	Black Dialect	MCM = 5[1] NMCM = 5[3] NMCM = 5[3] Total Miscues = 15[5]	MCM = 5[1] NMCM = 5[3] MCM = 5[4] Total Miscues = 15[5]

[1] Meaning-changing miscues: shared miscues, all versions.

[2] Non-meaning-changing miscues "typical" of Standard English speakers.

[3] Non-meaning-changing miscues reflection syntactic patterns of Black Dialect.

[4] Meaning-changing miscues "poor" reader version.

[5] Total miscues for 255-word passage.

version and poor reader version as read by the same adult. Individual reader speech features, other than those controlled through MCM and NMCM, were not controlled.

Procedure

The following scenario was established: subjects were told that the investigators were interested in exploring the factors that shape a person's diagnosis of a reader, and that by listening to tapes of readers, we were approximating a typical first step in classroom reading evaluation. Each subject was given four copies of the original story text and was told that (s)he would hear four children, each reading the material. Subjects were told to make whatever notes they felt necessary on the copies in order to rate and rank the readers.

The task for all the subjects was the same. On the first day of class, each subject listened to a tape recording of four female "children" reading the same 255-word selection. After hearing each reader, the subject completed a 10-item, five-point bi-polar adjective scale to rate the reader. Examples of dimensions on the scale are: the reader uses proper phrasing ... does not use proper phrasing; is confident ... is unsure; seems to read with meaning ... is a word caller. After a subject heard all four readers, (s)he was asked to rank order the four readers, with the best reader ranked first, and the poorest reader ranked fourth.

Results

The responses on the ten-item scale from each subject for each tape were summated, creating for each subject four scores, a summated rating of the SEGR, SEPR, BDGR and BEPR readers. In addition, each subject rank ordered the four readers (s)he heard. Table 2 contains the means and standard deviations of the summated rating scores and the ranking scores for all subjects for the four reader types.

Table 2

Means and Standard Deviations: Rating and Ranking Scores

	Ratings[1]		Rankings[2]	
	M	*SD*	*M*	*SD*
Standard English Good Reader	32.85	9.56	1.71	.939
Black Dialect Good Reader	27.56	6.94	2.28	.989
Standard English Poor Reader	23.21	7.78	2.96	.884
Black Dialect Poor Reader	23.45	7.86	3.03	1.11

[1] Rating scores could range from 10 to 50: summation of ten 5-point items.

[2] Ranking scores: range from 1 to 4 (with 1 being the most positive and 4 the least positive).

Rating and ranking scores were analyzed separately. Rating and ranking scores represent repeated measures on subjects and as such are correlated data. Both the rating data and the ranking data were, therefore, analyzed using a non-parametric Friedman ANOVA (Sigel, 1956). Significant differences were obtained between readers for both the ratings X^2 (3) = 57.44, $p<.05$, and the rankings, X^2 (3) = 58.76, $p<.05$. Post hoc comparisons of both the ratings and rankings by means of a nonparametric confidence interval procedure (Rosenthal and Ferguson, 1965) indicated similar patterns in the rating and ranking data; good readers as a group were rated and ranked more positively than poor readers. The SEPR and BDPR readers were not ranked differently from each other. However, the SEGR and BDGR readers were rated and ranked differently with the SEGR reader evaluated more positively in both instances.

A two-by-two multivariate analysis of variance was also done with the dependent variables being the four summated ratings of the four readers, and the independent variables being race of subject (Black vs. White) and whether or not a person had taught reading. No significant differences in rating behavior were found due to either of these factors or their interaction.

Conclusions

The results of this study suggest that teachers may confuse oral language features of Black Dialect speaking readers with reading ability. Black Dialect non-meaning-changing syntactic features embedded in the oral reading of otherwise good readers appear to have been salient in teachers' judgements of these readers. Support for this conclusion is found in both the rating scale data and the ranking data; Black Dialect speaking good readers were rated and ranked significantly lower than their Standard English speaking counterparts.

It is interesting that this finding did not hold for poor readers; BDPR and SEPR ratings and rankings were not significantly different. An interpretation of these differing results is possible if it is assumed subjects were "looking for mistakes."

The poor reader tapes contained more meaning-changing miscues than did the good reader tapes. The dialect-based miscues may have only affected judgements when these miscues co-occurred with a relatively small number of meaning-changing miscues. But when surrounded by a large number of meaning-changing miscues, the dialect-based miscues may be less obvious and their impact on judgements lessened.

These findings provide some empirical support for the hypothesis suggested by Goodman and Buck (1973) and Simons and Johnson (1974) that it is teacher response to language different children, not dialect difference per se, that may be related to many of these children's failure to learn to read. If teachers look for errors inthe oral reading of children and consider syntactic features of the oral language of nonstandard speaking children as errors, then these non-standard speaking children may be judged as poor readers. As Goodman and Buck suggest, such a judgement may result in teachers moving children away from their linguistic competence by requiring word-for-word accuracy.

There are several implications suggested by these findings. First, if teacher judgements about the reading abilities of nonstandard speaking children are a result of lack of knowledge of the features of nonstandard dialects, then teacher education must provide such information. Protocol materials by Love (1973) have been helpful in promoting such knowledge (Ramig, Granger & Neel, 1976).

However, if judgements about the reading ability of nonstandard speaking students are a result of a generalized stereotypic response to nonstandard dialects, then perhaps knowledge of the features of nonstandard dialects is not sufficient. Techniques for modifying such stereotypic expectations have been discussed by Billiard, Elifson and Rubadeau (1976).

Finally, if subjects' responses to the oral reading of children result from an attitude that only word-for-word precision is good reading, then study of psycholinguistic conceptualizations of the reading process is probably appropriate.

Teacher education ought to provide opportunities for teachers not only to learn about language and language differences as they relate to reading, but also must provide opportunities for teachers to develop human and positive attitudes toward language different children. Given appropriate learning, teachers may be able to distinguish reading ability from related but non-essential factors.

References

Baratz, J. C. Teaching reading in an urban Negro school system. In Joan C. Baratz and Roger W. Shuy (Eds.), *Teaching Black Children to Read.* Washington, D. C.: Center for Applied Linguistics, 1969.

Baratz, J. C. Relationship of Black English to reading: A review of the research. In James L. Laffey and Roger W. Shuy (Eds.), *Language differences: Do they interfere?* Newark, Delaware: International Reading Association, 1973.

Bereiter, C. & Engelmann, S. *Teaching disadvantaged children in the preschool.* New Jersey: Prentice-Hall, 1966.

Billiard, C. E., Elifson, J. M., & Rubadeau, J. W. kAttitudes about nonstandard dialects can be changed. In R. C. Granger and J. C. Young (Eds.), *Demythologizing the inner city child.* Washington: National Association for the Education of Young Children, 1976.

Covington, A. J. *A study of teachers' attitudes toward Black English: Effects on student achievement.* Unpublished doctoral dissertation, University of Pittsburgh, 1972.

Crowl, T. K. & MacGinitie, W. H. The influence of student's speech characteristics on teachers evaluation of oral answers. *Journal of Educational Psychology,* 1974, *66,* 304-308.

Clark, A. C. Space pet. In A. Harris, *MacMillan Reading Program* (4th year level). New York: MacMillan Co., 1966.

Cunningham, P. M. An inquiry into teachers' attitudes toward black-dialect-specific and non-dialect-specific reading miscues. In *Reading: Convention and inquiry.* Twenty-fourth Yearbook of the National Reading Conference, 1975.

Gess, L. *The effects of information which is provided to teachers concerning students on the attitudes and behaviors of the teachers and the students.* Unpublished doctoral dissertation, University of Michigan, 1969.

Goodman, K. S. Dialect barriers to reading comprehension. In Baratz, J. C. and Shuy, R. W. (Eds.), *Teaching Black Children to Read.* Washington D. C.: Center for Applied Linguistics, 1969.

Goodman, K. S. & Buck, C. Dialect barriers to reading comprehension revisited. *The Reading Teacher,* 1973, *27,* 6-12.

Goodman, Y. & Sims, R. Whose dialect for beginning readers? *Elementary English,* 1974, *51,* 837-841.

Granger, R. C., Mathews, M., Quay, L., & Verner, R. Teacher judgements of the communication effectiveness of children using different speech patterns. *Journal of Educational Psychology,* 1977, *69,* 793-796.

Guskin, J. T. *The social perception of language variation: Black and White teachers' attitudes towards speakers from different racial and social class backgrounds.* Unpublished doctoral dissertation, University of Michigan, 1970.

Hall, V. C. & Turner, R. R. The validity of the "different language explanation" for poor scholastic performance by black students. *Review of Educational Research,* 1974, *44,* 69-81.

Hockman, C. & Black, H. Dialect reading tests in the urban elementary school. *The Reading Teacher,* 1973, *26,* 581-583.

Holmes, M. B. *Interaction patterns as a source of error in teachers' evaluations of head start children* (Report No. OEC [illegible]52). New York: Associated YM-YWHAS of Greater New York, August, 1968.

McDavid, R. I. Dialectology and the teaching of reading. *The Reading Teacher,* 1964, *18,* 1964, 206-213.

Melmed, J. P. Black English phonology: The question of reading interference. In James L. Laffey and Roger W. Shuy (Eds.), *Language differences: Do they interfere?* Newward, Del.: International Reading Association, 1973.

Naremore, R. C. Teachers' judgements of childrens' speech: A factor analytic study of attitudes. *Speech Monographs,* 1971, *38,* 17-27.

Nolen, P. S. Reading nonstandard dialect materials. *Child Development,* 1972, *43,* 1092-1097.

Ramig, C., Granger, R. C. & Neel, J. *Graduate student/teacher perceptions of children's oral reading as a function of knowledge of dialect differences.* A paper presented at the Great Lakes Regional Conference of the International Reading Association, Milwaukee, February, 1976.

Rosenthal, I., & Ferguson, T. S. An asymptotically distribution free multiple comparison method with application to the problem of *n* rankings of *m* students. *The British Journal of Mathematical and Statistical Probability,* 1965, *18,* 243-254.

Sigel, S. *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill, 1965.

Simons, H. D. & Johnson, K. R. Black English syntax and reading interference. *Research in the Teaching of English,* 1974 *8,* 339-358.

Sims, R. *A psycholinguistic description of miscues generated by selected young readers during the oral reading of text material in Black dialect and standard English.* Unpublished doctoral dissertation, Wayne State University, 1972.

Steadman, J. & Adams, R. L. Teacher perception of behavioral adjustment as a function of linguistic ability in Mexican-American head start children. *Psychology in the Schools*, 1973, *10*, 221-225.

Whitehead, J. & Adams, R. L. Teacher perception of behavioral adjustment as a function of linguistic ability in Mexican-American head start children. *Psychology in the Schools*, 1973, *10*, 221-225.

Williams, F., Whitehead, J. L., & Miller, L. M. Ethnic stereotyping and judgements of children's speech. *Speech Monographs*, 1971, *38*, 166-170.

JOHN C. STANSELL
Texas A&M University

JEROME C. HARSTE
Indiana University

ROGER J. DE SANTI
University of New Orleans

The effects of differing materials on the reading process

The variability of the reading process across differing materials merits careful study, since understanding of the nature and degree of such variability has direct impact upon teaching strategies. Although discrete reading processes for differing materials have been proposed (Gibson & Levin, 1975; Robinson, 1975), basic research is warranted to describe the interaction of readers with such materials (MacGinitie, 1975-76).

There are five systems that cue meaning in reading: (1) cue systems within words, (2) cue systems within the flow of language at the sentence level (Smith, Goodman, & Meredith, 1970), (3) cue systems within the flow of connected discourse, including intersentence grammar, semantic redundancy beyond the sentence level, and organizational structure characteristic of various kinds of prose (Neuwirth, 1976; Meyer, 1976), (4) cue systems within the reader, including the experiential background and conceptual abilities, scripts, and/or schemata (Pearson & Nicholson, 1976; Anderson, Reynolds, Schallert, & Goetz, 1976; Schank, 1972), and (5) cue systems external to both language and the reader, such as pictures, charts, and graphs. It should be noted that not all cue systems are available at every step of the way.

Other factors may also affect cue system utilization. Barr (1975) and Harste and Burke (1976) have suggested that the focus of reading instruction greatly affects cue utilization and hence reader strategies for dealing with print. This effect was particularly noted for poor readers, with good readers seemingly outgrowing their instructional model.

Differences in literary and historical narrative may also be significant to the reader. Literary narrative may be described as more emotive in tone (Wheelwright, 1954) and more abstract (Moffett & Wagner, 1976) than historical narrative; the possible effects of these differences upon the reading process have not as yet been examined in research.

The present studies involved in-depth investigations of cue system utilization by three groups of readers: second-grade students, ninth-grade students, and mature adults past the age of sixty. Second-grade readers read materials which varied according to organizational structure, while ninth-grade and mature adult readers read literary and historical narratives. The studies yielded a view of striking similarities in cue system utilization across wide age variations.

Method

Samples

Study I. Six readers from one second-grade classroom were involved in this study. The teacher was asked to rank order all the children in her classroom. Children who had the same kindergarten and first grade teachers were identified within this ranking and a stratified random sample was drawn, consisting of two children with similar instructional histories from the top middle, and lower one-third of the ranking.

Study II. Six ninth-grade English students who were considered proficient readers were selected. The teacher was asked to identify and rank-order her top ten readers and the top six were chosen. Although it was not possible to precisely reconstruct the instructional histories of these readers, an interview concerning their perceptions of the reading process and preferred reading strategies showed that they generally shared the same set of views.

Study III. Four non-institutionalized persons, ages 62, 64, 65, and 82, who were rated as mature readers (Gray & Rogers, 1956) were selected for this study. Data from an initial interview were examined by a panel familiar with the Gray and Rogers (1956) scale, and this panel determined the reading maturity of the study participants. The same interview format employed in Study II was used to collect data concerning perceptions of reading and preferred reading strategies. This data indicated that the four persons held similar views and that these views tended to coincide with those of the readers in Study II.

Materials

Study I. Three selections were chosen for this study, all of which dealt with turtles (Scott Foresman, Level 6, 1971). The amount of available prior knowledge was therefore constant across the three selections. Stories were analyzed for readability using the Spache formula and were found to be highly similar at approximately 2.6. Materials were selected on the basis of readability and because they represented typical formats found in social studies materials, science materials, and literature. To verify selection choices, copies of all three selections were given to one class of undergraduate students in elementary education and each student was asked to classify each selection according to discipline. "Turtle Rescue" was identified by 83% of the students as a social studies selection, 100% identified "Kinds of Turtles" as a science selection, and 93% identified "Clever Turtle" as a literature selection.

Studies II and III. Two selections, both about the American Civil War, were selected for Studies II and III (Crane, 1952; Brown, Robinson, & Cunningham, 1974); thus, pertinent prior knowledge was held constant for the selections. The Dale-Chall and Flesch readability formulas were used to determine that the selections were comparable in terms of readability; both selections were rated as 7th to 9th grade material by both formulas. Both narratives employed a straightforward chronological pattern of organization. Neither provided extralinguistic cues. The significant difference in the narratives was taken to be that the literary selection is relatively more emotive and abstract than the historical selection, which presents a more nonemotive and concrete treatment of events.

Procedure

Study I. Procedures described in the Reading Miscue Inventory (RMI) Manual (Goodman & Burke, 1972) were used in collecting data from each of the six readers. The three selections were read by each reader in a single one-hour session, and the order of selection presentation was varied across readers. Reading sessions were conducted by a single researcher in a teacher's lounge.

Study II. Data collection procedures for this study are similar to those described for Study I. Since the two selections used in this study are considerably longer than those used in Study I, two reading sessions were held for each reader. In the intial session, the reader was interviewed, and read and retold one of the two selections. The order of selections read was alternated across readers. All reading sessions were conducted in a quiet room.

Study III. The procedures used in this study are identical to those described in Study II, except that all data-collection sessions were conducted by the researcher in the homes of the readers.

Analysis

Study I. Worksheets for each selection were prepared following procedures as described in the RMI Manual. Standard miscue counting procedures were followed with the exception that partial corrections were counted as miscues and coded as evidence of within-words cue utilization. The first 30 miscues made by a subject (or all miscues when less than 30) on each of the three selections were analyzed. Following this procedure produced 398 miscues for analysis. Twenty-four questions were asked of each miscue.

Miscues were coded in terms of the reader's utilization of each of the five cue systems. When there was evidence that cue systems within words had been utilized, the miscue was coded as partial word (letter-sound relationships or affixes) or whole known words (substitutions, omissions, or the production of known words within the text word on which the miscue occurred). Use of syntactic and semantic cues within the flow of language were coded according to whether these cues were at the sentence level or the discourse level. The reader's utilization of preferred oral language structures and lexical item preferences was taken as evidence of his bringing his conceptual understanding of language and the world to the task, and were coded within the category of cue systems within the reader. Extralinguistic cue system use was noted on the worksheet at the time of taping, and coded as pictures and other such cues as charts and graphs.

Ten additional questions coded which of the five cue systems had been used to predict and to confirm meaning.

Miscues were coded as to whether they resulted in meaning change. Meaning change was coded as *Yes, No*, or *Partial*, using decision rules established by Goodman and Burke (1976).

Percentages (proportion of actual cue system utilization and effectiveness to total possible cue system utilization and effectiveness) were calculated for each of the 24 questions, using the total number of miscues made on the selection as the base unit. This procedure permitted the analysis of miscues at the clause, sentence, and intersentence level and resulted in a minimal loss of information.

In addition to these analyses, one additional calculation was made. This was miscues per hundred words (MPHW), the ratio of miscues made to the total number of words in the selection X 100. This procedure equalizes miscue counts across selections of various lengths.

Studies II and III. Data analysis procedures were nearly identical for these two studies, and similar to those described for Study I. A total of 1,105 miscues were analyzed in Study II and 268 miscues were analyzed in Study III. Reading interview data were examined to permit comparison of described and actual reading performance on differing materials.

Comprehension was assessed by calculating retelling scores according to the guidelines in the *RMI Manual*, as well as by coding each miscue according to whether it resulted in meaning change, and if so, whether this change was minimal or substantial. To equalize miscue counts across selections of different lengths, the number of miscues per hundred words (MPHW) was calculated. Indices of the readers' tendency to recover from miscues which disturb meaning were also constructed in these studies. In Study II, residual miscues per hundred words (RMPHW — the ratio of miscues which disturb meaning to the total number of words in the selection X 100) was used as such an index. Study III used the percentage of residual miscue sentences (those sentences in which an uncorrected miscue caused meaning change) as such an index. Finally, in Study II the percentage of regressions made by each reader at the morphemic or sub-morphemic level was calculated in an effort to illustrate the tendency toward the use of cue systems within words.

Results and Discussion

The balance of this report will present those conclusions common to all three of the studies and samples from the data supporting them. Some important implications which arise from these studies will also be noted.

Consistency of Cue System Utilization

Readers in all three studies tended to use the same cue systems in a consistent fashion regardless of the nature of the materials read. Ellen, a second-grade subject in Study I (see Figure I), made extensive and consistent use of the within-words cue system on all three selections, relying much more heavily on whole words than partial words. Vince was the most consistent user of the various cue systems across the three selections (see Figure 2). These two subjects utilized discourse cues, cues within the reader, and extralinguistic cues much less frequently in the three selections. Ellen's use of picture cues in 10% of her miscues on "Turtle Rescue" represents the greatest use, for both these subjects, of information from these systems.

The pattern of consistent preference for word level cues across selections also appears in the data from Study II (see Table I). Debbie's miscues exemplify the overall results. For the entire sample, substitutions tended to bear high phonetic and graphic similarity to the text word on both of the selections; regressions on both selections also tended to occur at the morphemic and sub-morphemic levels. Relatively few miscues on either selection resulted in fully acceptable semantic structures and few indicated grammatical strength. Finally, the rate of miscue occurrence (MPHW) was similar for both selections.

Study III produced similar results from the same selections used in Study II. 50% of Blanche's substitution miscues (see Table 2) on both of the selections were highly similar to the text word in terms of both graphics and sound. Her rate of miscues per hundred words (MPHW), while far lower than that of subjects in Study II, was consistent across selections: 1.6 for the literary narrative and 2.1 for the historical narrative. All of the subjects in Study III produced a relatively high percentage of miscues on both selections (65% +) that were fully acceptable syntactically and semantically. This would indicate that while the older subjects were very consistent in cue system use across differing materials, their preferences for word cues were neither as strong nor as detrimental to the use of other cue systems as were such preferences among younger subjects.

Clear preference for word cues across samples and selections did appear to influence both the use of cues from other systems and comprehending effectiveness in all three studies. In the preceding data, it can be seen that the

40

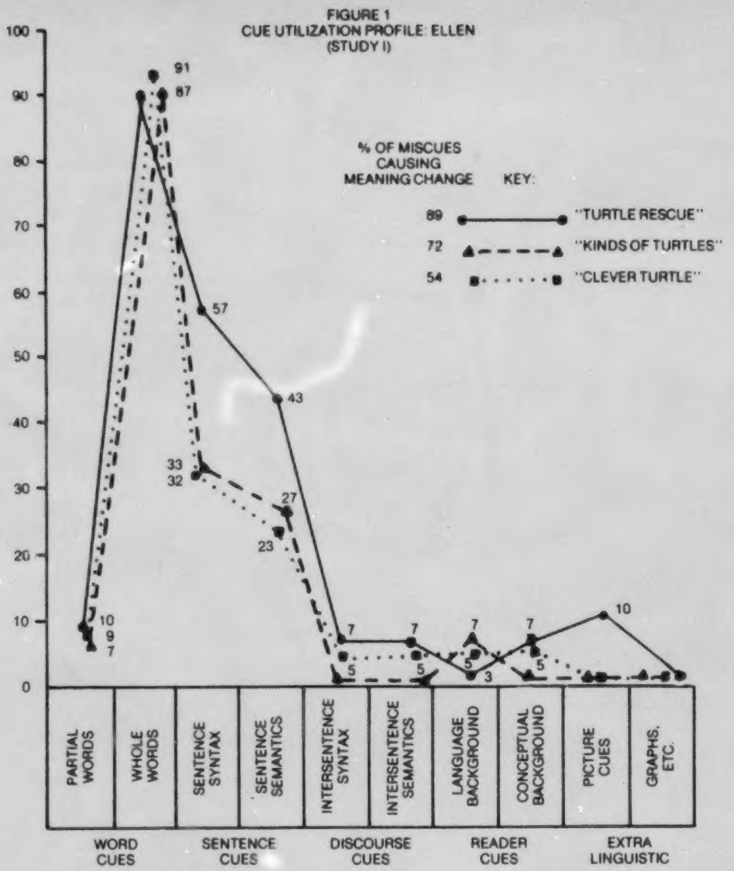


FIGURE 1
CUE UTILIZATION PROFILE: ELLEN
(STUDY I)

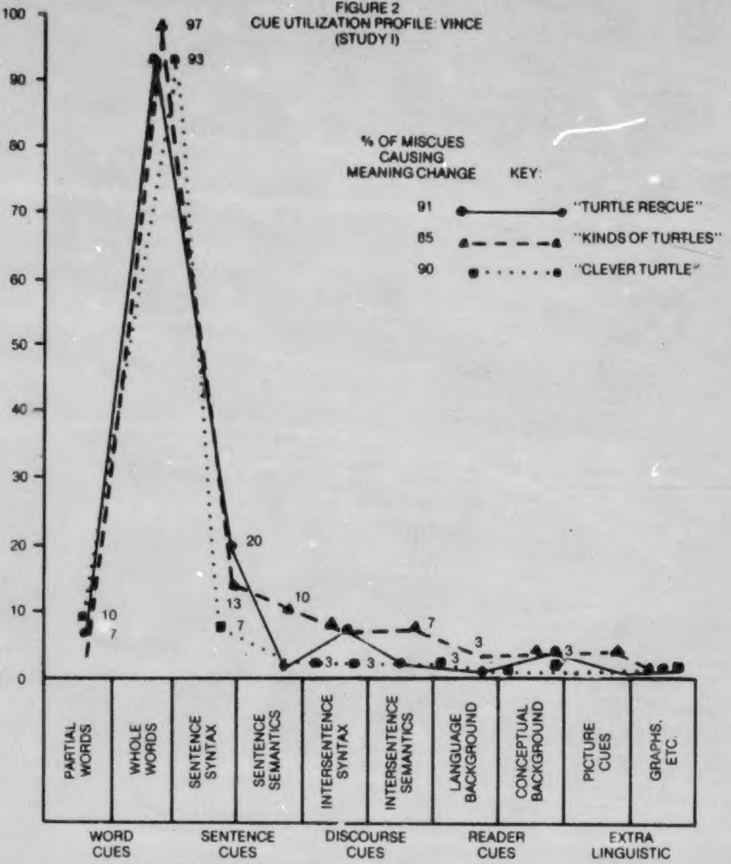
NOTE: ALL NUMBERS EXPRESSED AS PERCENTAGES

use of the word-level cue system is inversely related to the use of other available cue systems. The degree of preference for cues within words also seems to be inversely related to comprehending effectiveness.

Strong preference for cues within words on the part of Ellen and Vince (Figures I and II) was accompanied by high incidence of meaning change resulting from miscues. Debbie (Table I), who also consistently preferred within-words cues, was also a relatively poor comprehender of the two selections she read, as shown by a high percentage of miscues that caused meaning change, a high number of residual miscues per hundred words, and a low retelling score on both selections. Blanche's word cue preferences (Table II) were not as strong as those shown by younger readers. This is indicated by the low percentage of sentences on both selections in which she produced meaning - disturbing miscues, and by the relatively higher frequency of miscues that indicate grammatical strength. However, her retelling scores were quite low on both selections.

It should be noted that not all subjects in the three studies demonstrated an equally strong overreliance upon cues within words. The greatest variation in preferences was seen in Study I, where subjects who were rated as relatively proficient readers tended to be both more flexible in cue system use and

FIGURE 2
CUE UTILIZATION PROFILE: VINCE
(STUDY I)



NOTE: ALL NUMBERS EXPRESSED AS PERCENTAGES

better comprehenders. These readers did, however, demonstrate consistency in their patterns of cue system utilization.

Importance of the Reader's View of the Process

All subjects' reading performances tended to reflect their concepts of the nature of the reading process more clearly than the differences in materials. Vince and Ellen seemed to consistently view reading as a word-centered process regardless of the material they were reading, while other, more flexible readers in Study I were characterized by their strong and consistent reliance upon sentence and discourse level cues, and their conceptual and linguistic backgrounds. The interview data collected in Study II revealed that all subjects in that sample, like Vince and Ellen, tended to view reading as a process of accurate word recognition, and the oral reading data from this sample which was previously discussed indicates that reading performances were consistent, across both selections, with these stated perceptions of reading. Interviews among the older subjects of Study III revealed a significant concern for accuracy of word recognition, which was also apparent in their oral reading performances on the two selections. The reader's view of the process does seem to determine patterns of cue system use to a much

Table 1

Cue Utilization of Effectiveness in Study II

	Literary Narrative		Historical Narrative	
% of substitutions with high phonetic similarity	64	65	66	65
% of substitutions with high graphic similarity	67	68	71	69
% of total regressions occurring at morphemic and submorphemic levels	79	62	80	67
% of miscurs resulting in full semantic acceptability	15	26	32	40
% of miscues indicating strengths in preserving grammatical relationships	27	31	40	48
% of miscues resulting in meaning change	68	63	45	45
miscues per hundred words (MPHW)	15.12	6.94	13.47	6.39
residual miscues per hundred words (RMPHW)	10.85	4.19	7.28	2.76
retelling score	12	19	2	10

greater extent than the differences among materials from various content areas.

Implications

Data from these studies suggest that the reading process is relatively stable across content areas, and that the widely held notion of significant differences in the demands made by content materials (Piercey, 1976; Robinson, 1975; Gibson & Levin, 1975; Herber (1970) needs reexamination. However, strong concern for accurate decoding was related to poor comprehending, even among the high-school subjects and those past the age of sixty.

Whether or not the preference for attending to cues within words at the expense of meaning is an artifact of instruction is an important question that merits further study. The present studies, however, seem to clearly indicate a need for instruction which encourages readers to apply their knowledge of language and of the world in making greater use of all available cue systems, and avoiding overreliance upon those cues which, inflexibly used, lead to poor comprehension.

Table 2

Blanche's Cue System
Use and Effectiveness
(Study III)

	Literary Narrative	Historical Narrative
% of substitutions with high phonemic similarity	50	50
% of substitutions with high graphic similarity	50	50
% of miscues indicating strength in preserving grammatical relationships	76	57
% of sentences with miscues that disturb meaning	2	9
miscues per hundred words (MPHW)	1.6	2.1
retelling score	19	6

References

Anderson, R. C., Reynolds, R. E., Schallert, D. L., & Goetz, E. T. *Frameworks for comprehending discourse.* Urbana, Illinois: University of Illinois, Laboratory for Cognitive Studies in Education, Technical Report No. 12, 1976.

Barr, R. The effect of instruction on pupil reading strategies. *Reading Research Quarterly,* 1974-75, *10,* 555-582.

Crane, S. An episode of war. In R. W. Stallman (Ed.), *Stephen Crane: An omnibus.* New York: Alfred A. Knopf, 1952.

Gibson, E. J. & Levin, H. *The psychology of reading.* Cambridge, Massachusetts: MIT Press, 1975.

Goodman, Y. M. & Burke, C. L. *Reading miscue inventory manual: Procedures for diagnosis and evaluation.* New York: MacMillan, 1972.

Goodman, Y. M. & Burke, C. L. *The reading miscue inventory: Evaluation form.* Tucson, Arizona: Center for Expansion of Language and Thinking, 1976.

Gray, W. S. & Rogers, B. *Maturity in reading: Its nature and appraisal.* Chicago: The University of Chicago Press, 1956.

Harste, J. C. & Burke, C. L. A new hypothesis for reading teacher education research: Both the teaching and learning of reading are theoretically based. in P. David Pearson (Ed.), *Twenty-fifth Yearbook of the National Reading Conference*, 1976.

Herber, H. L. *Teaching reading in content areas.* Englewood Cliffs, New Jersey: Prentice-Hall, 1970.

MacGinitie, W. Research suggestions from the Literature Search. *Reading Research Quarterly*, 1975-76, *11*, 7-15.

Meyer, B. B. F. Organization in prose and memory: Research with application to reading comprehension. In P. David Pearson (Ed.), *Twenty-fifth Yearbook of the National Reading Conference*, 1976.

Moffett, J. & Wagner, B. J. *Student-centered language arts and reading, K-13* (2nd ed.). Boston: Houghton-Mifflin, 1976.

Neuwirth, S. E. A look at intersentence grammar. *The Reading Teacher*, 1976, *30*, 28-32.

Pearson, P. D. & Nicholson, T. *Scripts, texts, and questions.* Paper presented at the Annual Meeting of the National Reading Conference, Atlanta, Georgia, December 1976.

Piercey, D. *Reading activities in content areas.* Boston: Allyn and Bacon, 1976.

Robinson, H. A. *Teaching reading and study strategies: The content areas.* Boston: Allyn and Bacon, 1975.

Schank, R. C. Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, 1972, 3, 552-631.

Smith, E. B., Goodman, K. S. & Meredith, R. *Language and thinking in the elementary school.* New York: Holt, Rinehart and Winston, 1970.

Wheelwright, P. *The burning fountain.* Bloomington, Indiana: Indiana University Press, 1954.

The years of southern success. In Brown, R. C., Robinson, W. S. & Cunningham, J. T. (Eds.), *Man in America* (Contemporary Social Science Curriculum (CSSC) Series). Morristown, New Jersey: Silver Burdett, 1974.

JOHN E. READENCE
University of Georgia

R. SCOTT BALDWIN
The University of Tulsa

Effects of impulsivity-reflectivity and type of phonics instruction on reading achievement

Ausubel (1968) stated that no realistic system of teaching could afford to overlook individual differences and that such differences were expressed in, among other things, general mode of cognitive functions and approach to problem-solving situations. The impulsivity-reflectivity dimension of cognitive style describes one set of individual differences that affect thinking and problem-solving behavior in human beings.

Kagan, Rosman, Day, Albert and Phillips (1964) postulated that impulsivity-reflectivity relates to the selection and evaluation of solution alternatives in situations of high response uncertainty. When presented with an opportunity to consider alternative solutions to problems, an impulsive child tends to act upon his initial response with little reflection. Conversely, a reflective child delays before carrying out a solution hypothesis and actively considers each alternative.

Since the task of reading presents a situation of high response uncertainty, cognitive style may have particular relevance for reading. To be successful in reconstructing meaning from print, the reader must be a decision-maker who selects appropriate linguistic cues, considers prior experiences, and applies his language skill, all toward generating selective, differential hypotheses (Goodman, 1967). Studies comparing impulsives and reflectives on the variable of reading achievement (Kagan, 1965b; Johnson, 1969; Shapiro, 1974) have consistently demonstrated that reflectives perform better than impulsives on tasks involving word recognition, reading comprehension, and reading readiness. (For a complete review of the impulsivity-reflectivity dimension of cognitive style, see Readence & Searfoss, 1976.)

There are two hypothetical explanations for the apparent fact that mastery over the initial stages of reading acquisition is facilitated by reflective behavior, i.e., thoughtful analysis and a minimum of guessing:

H1: The nature of the reading process is such that children are at an advantage if they are reflective.

H2: Teachers and publishers organize reading instruction in ways which tend to be advantageous for reflective children.

If H1 is true, it suggests that being reflective is, with respect to reading, "better" than being impulsive and that if impulsive children are taught to be more reflective they should also become improved readers. If H2 is correct, then it implies that impulsives are merely "different" from reflectives rather than deficient in some sense. It also suggests that the acquisition of reading skills and strategies might be enhanced for both reflectives and impulsives through manipulations of the learning environment, including instructional materials.

Attempts at cognitive style modification (Debus, 1970; Denny, 1972; Jacobs, 1974; Stein, 1969) have been moderately successful in getting impulsive children to respond more reflectively; but these studies have been unable to demonstrate that modifying impulsivity has any positive effect on any achievement variable. Consequently, the available information appears to be unfavorable to H1. On the other hand, there has been virtually no research designed to test the hypothesis that using alternative instructional materials or methods will affect reading acquisition for impulsive and, or, reflective children.

The purpose of the present paper will be to consider the position that cognitive styles are sensitive to variations in instructional approaches in reading. This will be accomplished by reanalyzing, using difference scores, the original data from Readence and Baldwin (in press), an aptitude-treatment interaction study involving impulsivity-reflectivity and alternative phonics programs.

Original Analysis

Karlin (1975) described two general approaches to phonics instruction, synthetic and analytic. In a synthetic approach, readers are taught sound-symbol correspondences and are shown how to blend sounds and letters to form words; thus, instruction proceeds from part to whole. In an analytic approach, readers first learn familiar words and work with the sounds within them, and instruction proceeds from whole to part.

In the original study, vocabulary and reading comprehension scores of impulsives and reflectives were compared in schools primarily using either analytic or synthetic approaches to phonics. It was predicted that the disparity in achievement between the cognitive style groups would be greatest in schools using an analytic approach and least in schools employing synthetic phonics. The rationale behind this hypothesis was that while reflectives should fare well under either condition, intensive phonics instruction under the synthetic approach should be especially beneficial to impulsives since it would force attention to the kinds of graphic detail which these children are apparently inclined to ignore.

Method

Subjects. The subjects for the study were 260 second grade students from six elementary schools in a large Midwest city belonging to middle socio-economic levels. Each student had been exposed to only one basal reading program.

Instruments. The *Gates-McGinitie Reading Test* (GMRT), Primary Form B (1972) was used to measure achievement in vocabulary and comprehension. The *Matching Familiar Figures Test* (MFF), developed by Kagan (1965a), was used to measure impulsivity-reflectivity.

Procedure. Of the six schools, two had reading programs centered around a synthetic phonics approach while the other four were using an analytic approach. Four trained examiners administered the MFF and school personnel administered the GMRT. Students not classified as impulsive or reflective by the MFF were eliminated from the study. Eighty-nine students were found to be impulsive and 81, reflective. Thirty-nine impulsives received analytic phonics instruction and 50, synthetic instruction. Of the reflectives, 36 were taught by analytic phonics and 45 by synthetic instruction.

Design. The raw scores for each subject on the vocabulary and comprehension subtests of the GMRT were assigned to the appropriate cell in a 2x2 MANOVA design (Finn, 1974). Main effects were phonics approach and cognitive style. Vocabulary and comprehension were the dependent variables.

Results

There were no statistically significant differences in reading achievement between subjects taught by the analytic and synthetic approaches. However, reflectives exhibited significantly higher levels of achievement than impulsives, $F\ (2,165) = 3.55, p < .05$. Significant differences also emerged in the multivariate test for interaction between cognitive style and approach to phonics, $F\ (2,165) = 4.57, p < .05$.

Univariate analyses for each dependent variable revealed that the reflectives had significantly higher scores on both vocabulary ($M = 39.48$) and comprehension ($M = 26.30$) than the impulsives ($M = 37.21$ and 23.79), $F\ (1,166) = 4.44, p < .05$ and $F\ (1,166) = 7.14, p < .01$, respectively.

Since neither of the univariate interactions proved to be statistically significant, the multivariate interaction was explained through separate multivariate analyses which compared reflectives with impulsives in each of the phonics approaches. The results indicated that in the synthetic approach reflectives ($M = 41.33$) performed significantly better than impulsives ($M = 37.54$) only in vocabulary, $F\ (1,93) = 7.76, p < .01$. In the analytic approach, reflectives ($M = 25.72$) had significantly higher scores than impulsives ($M = 22.77$) only in comprehension, $F\ (1,73) = 3.98, p < .05$.

Discussion

The results seemed to support the notion that reflectives tend to be better readers than impulsives, and the multivariate interaction involving cognitive style and phonics approach was interpreted to mean that reflectives and impulsives are affected differently by various types of reading programs. However, the results provided no support for the hypothesis that impulsives would find intensive synthetic phonics instruction particularly beneficial.

New Analysis

The original analysis of the multivariate interaction resulted in a cumbersome set of statistics and no satisfactory means of visually representing the interaction. The present analysis is an attempt to circumvent those problems through the use of "different scores" in a procedure by Levin (1977). The analysis is based on the hypothesis that the achievement ratio involving vocabulary and comprehension in the original study was dependent upon cognitive style and instructional approach; that is, for a given cognitive style group in a given phonics approach, vocabulary scores might be relatively high while comprehension scores were relatively low; or the reverse could be true.

Design

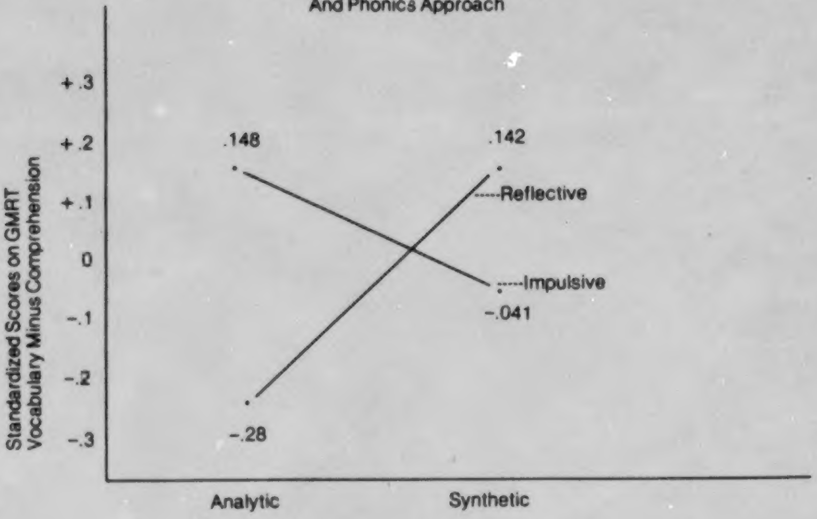
Vocabulary and comprehension subscores from the GMRT were transformed to Z scores. For each subject, the standardized comprehension score was subtracted from the standardized vocabulary score. The resulting difference score was then assigned to the appropriate cell in the 2x2 ANOVA design. Main effects were approach to phonics and cognitive style, and difference scores constituted the dependent measure.

Results

There were no significant differences between reflectives ($M = -.046$) and impulsives ($M = .042$), $F\ (1,166) = 1.00, p > .05$; and there were no significant differences between subjects in the analytic schools ($M = -.058$) and those taught by the synthetic approach ($M = .045$), $F(1,166) = .72$, $p > .05$. In contrast, there was a significant disordinal interaction, $F(1,166) = 8.82$, $p < .01$ (see Figure 1).

Figure 1

Plotted Cell Means Showing Disordinal Interaction Between Cognitive style And Phonics Approach



Discussion

The use of difference scores appears to be an efficient procedure for describing complex multivariate interactions. In this particular case, it clearly indicated that reflectives were relatively high achievers in reading comprehension under an analytic approach and relatively high in vocabulary when taught by a synthetic approach to phonics. Impulsives had a reversed achievement pattern. They were relatively high in vocabulary in an analytic program and high in comprehension under a synthetic approach.

The results are interpreted to mean that children with different cognitive styles are sensitive to different instructional approaches in reading. Furthermore, the authors find the results encouraging since it may be possible to factor out the characteristics of each instructional program which are uniquely beneficial to each cognitive style group. Once this is accomplished, individualized programs of instruction in reading could incorporate the best of each instructional approach in order to maximize reading achievement for both reflective and impulsive children.

References

Ausubel, D. P. *Educational psychology: A cognitive view.* New York: Holt, Rinehart, and Winston, 1968.

Debus, R. L. Effects of brief observation of model behavior on conceptual tempo of impulsive children. *Developmental Psychology*, 1970, *2*, 22-32.

Denny, D. R. Modeling effects upon conceptual style and cognitive tempo. *Child Development,* 1972, *43,* 105-119.

Finn, J. D. *A general model for multivariate analysis.* New York: Holt, Rinehart, and Winston, 1974.

Goodman, K. S. Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist,* 1967, *4,* 126-135.

Jacobs, P. C. Modifying impulsive behavior of second grade boys through cognitive self-instruction: Intensity of training related to generality and long range effects (Doctoral dissertation, Wayne State University, 1974). *Dissertation Abstracts International,* 1975, *35,* 4252-4253A. (University Microfilms No. 74-29, 816)

Johnson, B. L. Conceptual tempo and the achievement of elementary school boys (Doctoral dissertation, Case Western Reserve University, 1969). *Dissertation Abstracts International,* 1970, *30,* 3789A. (University Microfilms No. 70-4939)

Kagan, J. *Matching familiar figures test.* Unpublished test, Harvard University, 1965. (a).

Kagan, J. Reflection-impulsivity and reading ability in primary grade school children. *Child Development,* 1965, *36,* 609-628. (b).

Kagan, J., Rosman, B. L., Day, D., Albert, J. & Phillips, W. Information processing in the child: Significance of analytic and reflective attitudes. *Psychological Monographs,* 1964, *78,* (1, Whole No. 578).

Karlin, R. *Teaching elementary reading: Principles and strategies.* New York: Harcourt Brace Jovanovich, 1975.

Levin, J. R. Personal communication, University of Wisconsin, 1977.

Readence, J. E. & Searfoss, L. W. The impulsivity-reflectivity dimension of cognitive style. *Reading World,* 1976, *16,* 88-96.

Readence, J. E. & Baldwin, R. S. The relationship of cognitive style and phonics instruction. *Journal of Educational Research,* in press.

Shapiro, J. E. *The relationship of reading readiness skills to reflection-impulsivity and to the effects of visual discrimination training on impulsive first grade boyds.* Unpublished doctoral dissertation, Syracuse University, 1974.

Stein, N. L. The modification of impulsivity using two types of cue discrimination training (Doctoral dissertation, Stanford University, 1969). *Dissertation Abstracts International,* 1970, *30,* 5303A-5304A. (University Microfilms No. 70-10,532)

JEROME A. NILES
BARBARA M. TAYLOR
Virginia Polytechnic Institute and State University

The development of orthographic sensitivity during the school year by primary grade children

A number of theories of reading explain that knowledge or orthographic structure is used by readers to help them process text efficiently (Gibson and Levin, 1975; Smith, 1971). Although orthographic sensitivity is regarded as important in the reading process, there is disagreement as to when this sensitivity develops.

Smith (1971) feels that the ability to make use of orthographic structure develops as early as first grade. Studies by Lott and Smith (1971), Niles (1976), and Niles, Grunder and Wimmer (1977) support this viewpoint. Lott and Smith (1970) reported that first grade children were able to recognize letters at lower intensities when the letters were in words instead of in isolation. Niles (1976) found that first grade children used less visual information to recognize word-like pseudowords than non-wordlike pseudowords. Niles, Grunder, and Wimmer (1977) found that end of the year first grade children performed above chance on a task which involved selecting the more word-like pseudoword out of pairs of zero-order and fourth-order pseudowords. Kindergarten children performed at chance level on this task. From these studies it appears that orthographic sensitivity begins to develop at the earliest stage of learning to read.

Gibson and Levin (1975) support a different viewpoint, however, as to when orthographic sensitivity develops. They conclude from a number of studies that the sensitivity to orthographic structure begins to develop later than first grade. Gibson, Pick, and Osser (1963) found that first grade children did not read and spell out pronounceable four- or five-letter pseudowords any better than unpronounceable ones. Rosinski and Wheeler (1972) reported that first grade children performed at chance level on a task of selecting the more word-like pseudoword from pairs of pronounceable and unpronounceable pseudowords. Third grade children, in contrast, performed at better than chance. In a similar study Golinkoff (1974) found that students at the end of first grade performed only slightly better than chance on a task of selecting the more word-like pseudoword from pairs of visually presented pronounceable and unpronounceable pseudowords. Students at the end of second grade performed at a level significantly better than chance on this task. These studies suggest that first grade children are not sensitive to orthographic structure but that students at the end of second grade or in third grade have developed orthographic sensitivity.

The studies cited above present an unclear picture of the development of sensitivity to orthographic structure. Because the studies were based on different cross-sections of students and tested for orthographic sensitivity on different tasks at different times of the year, it is not surprising that there are conflicting findings. An extensive study which is longitudinal as well as cross-sectional should provide a more complete and accurate description of the development of orthographic sensitivity in beginning readers. Therefore, in the present study beginning readers were examined over time on a task measuring orthographic structure, an important text processing strategy, in the early stages of reading.

Additionally, this study will attempt to examine the relationship between reading ability and sensitivity to orthographic structure.

Method

Twenty first, second, and third grade children were randomly selected from each of the three grades in a rural elementary school. There were about 35 students in each grade at this school. Ten sets of eight-letter approximations to English (Miller, Bruner, and Postman, 1954), zero-order (hhjhufsw), second-order (riprypli), and fourth-order (inforems), were presented individually to the subjects in the second half of October, May, and the following October. Attrition over the one-year reduced the number in each group to 17. Each subject was tested individually and asked to select the one pseudoword from a set of three (zero, second and fourth order) which looked most like a real word. After making this first selection, the subject was then asked to select the pseudoword from the remaining two in the set which looked most like a real word. The set members as well as the ten sets were presented in random order. The ranking responses were scored using a four point scale for the six possible rankings. A chance score for the task was considered to be 25 and a ceiling score obtained from a group of college-educated adults was 39.37. At the time of the final testing the subjects were administered the reading comprehension subtest of the *Peabody Individual Acnievement test* (Dunn and Markwardt, 1970). A 3 X 3 (grade X time) factorial design was used with repeated measures on the second factor. Data were formally analyzed using an analysis of variance procedure and Newman Keuls Test for significant mean differences.

Results

Table 1 presents a summary of the means and standard deviations for grades 1, 2, and 3 by the specific time of the measurement. An examination of

Table 1

Means and Standard Deviations of Ranking Responses by Grade and Time

Time		Grade		
		1	2	3
Fall 1	*M*	26.19	31.38	33.5
	SD	3.59	4.53	3.09
Spring	*M*	27.94	34.19	34.31
	SD	4.0	3.28	4.28
Fall 2	*M*	31.0	34.94	35.88
	SD	3.75	2.96	3.86
Total	*M*	28.38	33.5	34.56

the means within each grade level suggests a developmental trend in the acquisition of sensitivity to orthographic structure. The age-related nature of orthographic knowledge is reflected in the overall means for each grade, 28.38 for first, 33.5 for second, and 34.56 for third. A formal analysis of this data revealed no significant interaction between variables and a significant main effect for both grade level and time $F(2,45) = 18.87$, $p < .01$ and $F(2,45) = 19.93$, $p < .01$, respectively.

The Newman-Keuls test revealed a number of significant differences among means. All significant differences were $p < .05$. While the within grade means increased on each subsequent measurement, only the first grade means all differed significantly from each other. Second grade means differed significantly from Fall 1 to Spring but not from Spring to Fall 2 measurement. Across grades, the first grade means were significantly different from all other means in every case, with the exception of Fall 2 to second grade Fall 1. Fall 1 for second grade was the only second grade mean to reflect a difference when compared to the third grade scores.

A chance score on the task was 25. A t test confirmed that the Fall 1 mean for first grade did not differ significantly from chance, ($p > .05$). However, the mean for the Spring of first grade did reflect a significant difference when compared with the chance score, ($p < .01$).

The Pearson correlations between reading ability (the comprehension subtest of the *Peabody Individual Achievement Test*) and the orthographic structure task were .66, .49, and .60 for the Fall 2 scores of the first, second, and third grades respectively and .68 for all grades.

Discussion

These findings clearly support the notion that, in general, sensitivity to orthographic structure is acquired by readers during the latter part of first grade even though the differences among cell means suggest that the growth of knowledge for orthographic structure was rapid throughout first grade. All conditions for the first grade were significantly different from each other. The second grade children demonstrated significant growth from the fall measurement to the spring measurement. At this point acquisition apparently slowed as there was no significant difference between the end of the year second grade children and the final test on them when they were beginning third grade. The acquisition rate slowed even more for the third grade children as they showed no significant growth from the Fall 1 to the Spring measurement. However, the third grade children did reflect significant growth when the Fall 1 and Fall 2 scores were compared.

This growth pattern quite obviously argues against the notion that sensitivity to orthographic structure does not begin to develop until the end of the second grade. In fact, for this task, the growth rate for the acuqisition of orthographic knowledge begins to slow down toward the end of second and beginning of third grade. It would seem that with the onset of learning to read, which generally occurs in first grade, the reader begins to rapidly develop orthographic knowledge. As a child begins to develop fluency in reading toward the end of the third grade, the acquisition curve for orthographic sensitivity, at least for this task, begins to gradually move toward an adult performance level.

Most likely, the marked difference in findings between this study and others that document the onset of sensitivity to orthographic structure at a later point is related to the type of stimuli used. While eight-letter pseudowords were used in the present task, shorter stimuli were used in other investigations.

For example, Rosinski and Wheeler, (1972) used three, four, five and six letter pseudowords and Gibson, et.al., (1963) used four and five letter pseudowords. It is quite possible that the stimuli in these studies were not sensitive enough for the first and second grade students to demonstrate their orthographic knowledge. The longer stimuli may have provided the additional redundancy which enabled the first and second students to make more accurate orthographic decisions.

The relationship between reading ability and the orthographic task also supports the findings of early development of sensitivity to orthographic structure. If reading ability and the task are correlated, it seems that one might logically expect to find evidence of orthographic knowledge at almost any point at which reading ability is discernable. While beginning readers may not be as knowledgeable about orthography as fluent readers, it is unreasonable to assume that this knowledge base is as slow in developing as other researchers have suggested. Knowledge of the orthography helps make reading easier and the inherent tendency of the human information processor seems to be to make the perceptual process as efficient as possible.

References

Dunn, L. & Markwardt, F. *Peabody Individual Achievement Test.* Circle Pines, Minnesota: American Guidance Service, Inc., 1970.

Gibson, E. & Levin, H. *The Psychology of Reading.* Cambridge, Mass.: MIT Press, 1975.

Gibson, E. J., Osser, H. & Pick, A. A study in the development of grapheme-phoneme correspondences. *Journal of Verbal Learning Verbal Behavior,* 1963, *2,* 142-146.

Golinkoff, R. *Children's discrimination of English spelling patterns with redundant auditory information.* Paper presented to American Educational Research Association, February, 1974.

Lott, D. & Smith, F. Knowledge of intraword redundancy by beginning readers. *Psychonomic Science,* 1070, *19,* 343-344.

Miller, G. A., Bruner, J. S. & Postman, L. Familiarity of letter sequences and tachistoscopic identification. *Journal of General Psychology,* 1954, *50,* 129-139.

Niles, J. The use of featural and letter dependency information in word recognition by elementary school children. *Reading Research Quarterly,* 1976, *II,* 198-202.

Niles, J., Grunder, A. & Wimmer, C. The effects of grade level and setting on the development of sensitivity to orthographic structure. In P. D. Pearson (Ed.), *Reading: Theory, research and practice,* Twenty-sixth Yearbook of the National Reading Conference. Clemson, South Carolina: 1977.

Rosinski, R. & Wheeler, K. Children's use of orthographic structure in word structure in word discrimination. *Psychonomic Science,* 1972, *26,* 97-98.

Smith, F. *Comprehension and learning.* New York: Holt, Rinehart and Winston, 1975.

Smith, F. *Understanding reading.* New York: Holt, Rinehart and Winston, 1971.

ELIZABETH ARNOLD
GEORGE MC NINCH
WALLACE MILLER
University of Southern Mississippi Reading Center

Experiments in word learning

Many significant models of beginning and mature reading or their featural components appear in the recent literature (Geyer, 1970; Gibson & Levin, 1975; Guthrie, Goldberg, & Finacci, 1972; Singer, 1970). However, while the models are generally widely discussed in professional journals and conferences, they are not usually widely researched. Often it appears that only the primary authors or their closest students do a systematic evaluation of the proposed model. This is not an indictment, but rather a conclusion that led to the formation of the present study.

Gibson and Levin (1975) define a word as a "complex of features, a composite representation of five classes of information: graphic, phonological, orthographic, semantic, and syntactic" (p. 194). These features, according to Gibson and Levin, may or may not be extracted in the perceptual process of reading.

The problem, therefore, is to determine on what basis beginning readers learn to recognize and respond appropriately to the printed word, and to identify which, if any of the five classes of information are useful as pronunciation cues to naive beginning readers. Specifically, the purposes of the study are to investigate the following questions:

1. Do naive beginning readers respond differentially in the graphic domain to the visual memory of words for the graphic features of lower- and upper-case script?
2. Do naive beginning readers respond differentially in the phonological domain to phonologically pronounceable words and phonologically unpronounceable words?
3. Do naive beginning readers respond differentially in the orthographical domain to orthographically legal words and orthographically illegal words?
4. Do naive beginning readers respond differentially in the semantic domain to concrete and abstract words?
5. Do naive beginning readers respond differentially in the syntactic domain to nouns and verbs?

Sample

The sample of 20 students was selected from Gaston Point Elementary School and West Ward Elementary School in the Gulfport, Mississippi, Municipal Separate School District (*n* = 790). Both schools are fully integrated on a community basis; however, Gaston Point is predominantly Negro and West Ward is predominantly Caucasian.

To obtain the experimental sample, 40 students who obtained average or above readiness ratings on the school-given Metropolitan Readiness Tests, were randomly chosen from each of the schools. To further define the functioning levels the 40 students were administered the Lorge-Thorndike

Cognitive Abilities Test. Finally, the results of the two screening tests were collapsed and 10 students having both average or above readiness and intelligence levels were randomly selected from each school for the final experimental number of 20. Accordingly, the students could be predicted as possible successful learners based upon the two screening tests.

Experiments

Experiment 1

Experiment 1 was designed to investigate whether beginning readers respond differentially in the graphic domain to the visual memory of words for the graphic features of lower- and upper-case script.

The task consisted of a delayed matching-to-sample procedure. The stimuli were five, three-letter and five, four-letter real or legal nonsense words.

The words were selected by choosing a key word (stimulus), then changing the first letter for one response choice, changing the last letter for one response choice, and presenting the same word as the remaining choice (for example, MAP: MAT, MAP, BAP).

In order to distribute possible practice effects across the learning trials, children were randomly assigned to lists and rotated by receiving upper case and lower case and vice versa over a period of two days. A single "word" printed in primary type was presented on a Keystone Tac-ette tachistoscope for a one second exposure. Then an array of three randomly arranged response words printed in primary type was presented on a 5" x 5" card to the subjects. There were three practice items for both lists.

The mean numbers of correct responses for each of the two lists (lower and upper case) of Experiment 1 are presented in Table 1. Since each student served as his own control, a *t* test for the significance of the difference between correlated means for the lower- and upper-case lists was conducted. The first hypothesis was accepted on the basis of the finding and there appeared to be a difference in memory of word form with lower-case letters being recalled more frequently than upper-case letter, $t(19) = 2.39, p < .05$.

Experiment 2

Experiment 2 was designed to investigate whether beginning readers respond differentially in the phonological domain to phonologically pronounceable words and phonologically unpronounceable words.

The 20 children in the sample were taught by a prompting technique the sounds of five letters in isolation (hard *g*, *a*, *b*, *t*, and *n*). After the children were able to produce each sound when the letter was presented to them on a 5" x 5" card, the test items were administered by the examiner.

There were 10 items in the task consisting of 3 words, 2 of which were unpronounceable. All of the words were composed of 3 of the letters which had been previously taught to the children. Throughout the 10 items 1 pronounceable and 2 unpronounceable variants of the 3 letters were randomly arranged on 5" x 5" cards (for example, atb, tba, tab).

Each student was seated opposite the examiner. S/he was told that s/he would be shown three groups of letters and was instructed to point to the one he could say or pronounce.

Raw scores consisted of the number of identifications of pronounceable words. The mean ($M = 4.4$) which was computed for correct responses is shown in Table 1. An independent *t* test was utilized to determine whether the mean number of correct responses was significantly greater than chance

Table 1

Means, Standard Deviations, and *t*-test Values

for the Variables of Experiments 1-5

Variable	*M*	*SD*	*t* Value	*df*	*p*
			Experiment 1		
Upper	5.15	1.93	2.39	19	$p < .05$
Lower	6.40	2.30			
			Experiment 2		
Pro-nounceable	4.40	2.37	2.08	19	$p > .05$
Chance	3.33	2.37			
			Experiment 3		
Ortho-graphically Correct Words	4.10	1.83	1.95	19	$p > .05$
Chance	3.33				
			Experiment 4		
Concrete	6.10	4.61	.50	19	$p > .05$
Abstract	6.70	4.38			
			Experiment 5		
Nouns	2.55	1.57	2.91	19	$p < .05$
Verbs	6.15	5.41			

expectancy ($M = 3.33$). On the basis of the findings of the analysis, $t\,(19) = 2.08$, $p > .05$, the research hypothesis was not accepted and it was determined that there was no significant difference between the recognition of phonologically pronounceable words and phonologically unpronounceable words. Phonology when measured by reactions to pronounceable and unpronounceable words does not appear to be a task related to the recognition of words by naive beginning readers.

Experiment 3

Experiment 3 was designed to investigate whether beginning readers respond differentially in the orthographic domain to orthographically legal and orthographically illegal words.

A matching task was devised in which an abstract picture was presented along with three nonsense words. Two of the nonsense words had illegal orthography and one had legal orthography (for example, kcagl, ckalg, glack).

There were 10 items with nonsense words of 4, 5, and 6 letters in length with the legal word being presented in a random position in the list of 3 words. The words were typed in primary type and presented along with a picture on 5" x 8" index cards.

Each student was seated opposite the examiner and was told to point to the word that "tells what the picture is." S/he was instructed to guess if s/he was not sure. After each response the examiner said "okay" or "all right."

The mean number of correct responses ($M = 4.1$) was tested for significance by an independent t test to determine whether it was significantly greater than chance expectancy ($M = 3.3$). Results of the t test are shown in Table 1. On the basis of these findings, $t\ (19) = 1.95$, $p > .05$, the research hypothesis was not accepted and it was ascertained that there was no significant difference between the recognition of orthographically legal and illegal words by beginning readers.

Experiment 4

Experiment 4 was designed to investigate whether beginning readers respond differentially in the semantic domain to concrete and abstract words.

A criterion learning task was devised using six, four-letter words (three concrete: calf, park, dust, and three abstract: with, good, luck) selected from the Dale List of 769 Easy Words according to the following factors: (a) word length, (b) configurational elements, (c) different initial and final letter, (d) freedom from obvious discrimination confusion, and (e) words not used in the beginning basal material in the cooperating school system.

A flash device was used to produce the learning condition. The device was made from a desk telephone number index that allows for the exposure of the stimuli words in a predetermined random order by sets (concrete or abstract) to provide a systematic randomization in order to spread learning effect across trials.

After determining that the students did not know the words (Trial 1), each child was individually taught the words by a prompting technique until all of the words were known as evidenced by a trial of three correctly pronounced words. The number of trials to criterion for each set of three words was the raw score used for the statistical analysis. Task words by sets were learned at one day intervals with random assignment to the list to be learned first.

The raw scores used for analysis consisted of the number of trials to criterion required by each individual for each of the classes of words (concrete and abstract). The mean number of trials required by the subjects for the mastery of each of the lists of words is presented in Table 1. Since each subject served as his own control, a t test for the significance between correlated means was conducted. Although the mean number of trials for the abstract list ($M = 6.7$) was greater than the mean number of trials for the concrete list ($M = 6.1$), no statistically significant difference was found in the analysis between the two groups, $t\ (19) = .50$, $p > .05$. It, therefore, appears that concreteness and abstractness are not determiners of word reading facility under a testing/teaching condition with naive beginning readers.

Experiment 5

Experiment 5 was designed to investigate whether beginning readers respond differentially in the syntactic domain to nouns and verbs.

As in Experiment 4 a word learning task was devised in which six common words (three nouns and three verbs) were selected according to the following factors: (a) word length, (b) configurational elements, (c) different initial and final letters, (d) freedom from obvious discrimination confusion, and (e) words not used in the beginning basal materials in the cooperating system.

The desk telephone index (described in Experiment 4) was again utilized to produce the learning condition.

The experiment, as in Experiment 4, was a trials-to-criterion learning task. The raw scores used for analysis (as in Hypothesis 4) consisted of the number of trials to criterion required by each individual student for both classes of words (nouns and verbs). The mean number of trials required by the students for the mastery of each list of words is presented in Table 1. Hypothesis 5 was accepted on the basis of the finding of the analysis, in that the students required significantly more trials to learn the verbs than the nouns, $t(19) = 2.91, p < .05$. Form class may, therefore, be a determiner of word learning facility under a perception/learning technique with young learners.

Summary

In regard to the five questions posited for investigation the following significant results were noted.

Significant differences were found for lower- and upper-case script with lower-case letters being more easily procesed in short term memory than upper-case letters and for the response alternatives with initial letters being more salient cues than final letters. A significant difference was also found between the learning of nouns and verbs with subjects learning nouns more easily than verbs. No significant differences were found between the recognition of pronounceable and unpronounceable words and orthographically legal and illegal words. Also, no significant differences were found between the learning of concrete and abstract words and learning words with or without pictures.

It was concluded that the model of word information cues as posited by Gibson and Levin (1975) does seem to be reflected, at least to some extent, in the word perception skills of naive beginning readers. In the graphic domain, it was concluded that configuration as formed by the ascending and descending letters of lower-case script is a cue for the visual memory of words and that initial letters are more salient cues than final letters. In the phonological and orthographic domains, the authors conjectured that even though information was not extracted by the subjects, the skills to extract phonological and orthographic information might develop with an increase in age and reading ability. In the semantic domain, it was concluded that neither concreteness nor abstractness seems to affect word learning. In the syntactic domain, it was concluded that nouns are more easily learned than verbs. Also, the addition of picture cues to a word learning task does not appear to influence learning efficiency.

References

Geyer, J. J. Models of perceptual processes in reading. In H. Singer & R. B. Rudell (Eds.), *Theoretical models and processes of reading*. Newark, Delaware: International Reading Association, 1970.

Gibson, E. J. & Levin, H. *The psychology of reading.* Cambridge: MIT Press, 1975.

Guthrie, T. J., Goldberg, H. R. & Finucci, J. Independence of abilities in disabled readers. *Journal of Reading Behavior*, 1972, *4*, 129-138.

Singer, J. Theoretical models of reading: Implications for teaching and research. In H. Singer & R. B. Rudell (Eds.), *Theoretical models and processes of reading.* Newark, Delaware: International Reading Association, 1970.

60

ELIZABETH SULZBY*
Northwestern University

Children's explanations of word similarities in relation to word knownness

Researchers have been gathering evidence over the past few years that young children beginning to learn to read do not have a very firm grasp on the meanings of linguistic terms which are used in school instruction; they do not use *meta*language, or language referring to linguistic units and functions, very accurately (Downing & Oliver, 1974; Ehri, 1975, 1976, 1977; Holden & MacGinitie, 1971; Mickish, 1974; Karpova, 1955; Papandropoulou & Sinclair, 1974; Tovey, 1976).

One important metalinguistic concept for beginning readers is that of "word". Is a word to a beginning reader primarily a structural unit of letters and/or sounds? Or is it a unit which can enter into syntactic arrangements? Or is it primarily a unit to which people assign meanings, a semantic unit? Do children respond differently to written words than to oral words? Do they respond differently depending upon whether or not the word is known to them in written form? Is it true that, once words are learned in written language, it is possible to direct children to attend to different aspects of words, such as the phonology or the orthography of words, the syntactic arrangements words can enter into, or the semantic functions words can play (Gibson, 1971; Mason, 1975a, 1976b)? A number of researchers (Golinkoff & Rosinski, 1976; Rosinski, Golinkoff, & Kukish, 1976; Wickens, 1970, 1972) have contended that, while children can be directed to attend to phonological and syntactic aspects of words, semantic aspects take precedence in the child's developing concept of "word".

The present study investigated the child's development of the concept that a word is a unit used to express meanings (it also investigated the metalanguage used by the children during the experimental task). Pilot studies had indicated that younger children tend to think of words in written language primarily as structural units of letters and/or sounds and that older children think of words in written language primarily as semantic units.

In the pilot studies, however, and in previous studies into the development of semantic notions (Anglin, 1970; Naron, in press), the same words were used for all subjects. The pilot studies, design and measurement theory (Clark, 1973), developmental theory (Brown, 1975), and developmental research (Richman, Nida, & Pittman, 1976) indicated that the sample of linguistic units should be as carefully selected as the sample of human subjects; the sample of linguistic units (here, of words) should take into account the subject's prior knowledge of the units. For these reasons, it was considered necessary to identify and use words which the children in the study rated as known and unknown in written language.

---

* The author is indebted to the following people for their help in making this study possible: Thomas Estes, James Deese, Edmund Henderson, & Herbert Richards, all of the University of Virginia, and the administration, faculty, and children of Sussex County, Virginia.

Method

Subjects

Thirty children each from grades one, two, four, and six were chosen at random from an elementary school in a rural Virginia county. In this school system no one method of teaching reading was used to the exclusion of other methods, and children were placed in classrooms using chronological age as the primary criterion under a continuous progress model. Achievement and ability scores in the county averaged below national norms; however, the range of scores was equivalent to the national range. The racial breakdown of students (75% black and 25% white) was within 5% of the population totals for the county.

Procedure and Design

Each child took part in two tasks: A Knownness Rating in which his or her words were selected and a Word Choice task in which the child explained how chosen words went together for him or her. The words for the Knownness Rating came from word lists of nouns chosen at random from four sources: *Harris & Jacobson core words for first grade* (Harris & Jacobson, 1973); *Sheldon basic reading series* (1968); *Clarence R. Stone's revision of the Dale list of 769 easy words* (Spache, 1968); and *The teacher's word book of 30,000 words* (Thorndike & Lorge, 1944). *The American heritage dictionary of the English language* (1975) was used to determine if the most common usage of the randomly selected word was indeed noun. Words were arranged in lists according to frequency of occurrence to aid the children in selection. From the lists each child rated 24 words as known and 24 words as unknown to him/her in written form. The Knownness Ratings were conducted a day prior to the Word Choice task.

In the Knownness Rating examiners scored a word Known (K) if the child pronounced the word correctly (within the child's speech patterns) and if the child gave a gist of meaning for it appropriate to the community. The 48 words from the Knownness Rating were cast into four knownness configurations for the Word Choice task. The four knownness configurations were used to study the effects of the child's knowing all of the words (K-KK); knowing none of the words (U-UU); knowing either his/her word, the question word (K-UU); or knowing only the examiner's words, the choice words (U-KK), in an X:XA,XB paradigm. The four knownness configurations (K-KK, K-UU, U-KK, U-UU) and the two presentation modes (written and oral) were counterbalanced and assigned to subjects at random in the Word Choice task.

The dependent measure was the children's explanations of choices in the Word Choice task. Children were asked to explain: "How does your word [the question word] go with my word [the choice word] for you?" Prior to the experimental task, the examiner used a sample triad, eliciting or modelling answers that were both semantic and structural in nature, high level as well as trivial. Responses and probes were recorded on individual protocols. Content analysis indicated that the intention of the responses could be reliably scored on a 3-point scale (1 = structural responses; 2 = unclassifiable responses; 3 = semantic responses).

Independent variables were grade level, knownness of the question word, knownness of the choice word, and mode of presentation. Protocols were used for descriptive analysis of metalanguage, as well as for the quantitative analyses.

Results

In a four-way repeated measures analysis of variance children were found to give significantly more semantic responses in three conditions: 1) if the question word was known as opposed to unknown, $F(1,119) = 36.48, p < .01$, means 9.0 and 8.4, 2) if the choice word was known $F(1,119) = 56.17, p < .01$, means 9.1 and 8.3, and if the word was oral rather than written, $F(1,119) = 31.77, p < .01$, means 9.1 and 8.3. Contrary to the results of the pilot studies in which knownness was not controlled, main effect for grade level was not significant. Interactions were significant for knownness of choice word and mode of presentation, $F(1,119) = 23.08, p < .01$, and for grade, knowness of the choice word, and mode of presentation, $F(3,116) = 3.67, p < .01$. Two-way analyses of variance treating each knownness configuration separately again revealed no main effects for grade. Significant main effects were present for mode of presentation in each of the three configurations in which at least one word was unknown (K-UU); $F(1,116) = 29.01, p < .01$; U-KK: $F(1,116) = 10.66, p < .01$; U-UU: $F(1,116) = 25.05, p < .01$). Interactions between mode of presentation and grade were present in the K-UU configuration, $F(3,116) = 3.69, p < .05$, and the U-UU configuration, $F(3,116) = 2.68, p < .05$. Examination of the interactions indicated that the youngest children treated words unknown in written language more often as structural units and words in oral language more often as semantic units. By sixth grade, however, all children were responding to words unknown in written language primarily as semantic units. Because the interactions here involve assuming that the child knows the words when presented orally, one-way analyses of variance for each configuration were conducted to see if the increase in semantic responses for words in written language was significant; results were not significant. (It should be noted that children's explanations were scored according to the judged intention of the explanation. "I can't answer because I don't know the meaning of that word," would be a semantic response.) The two-way analysis of the K-KK configuration indicated that children of all ages treated words known in written language semantically as frequently as they did words in oral language.

Descriptive analyses of the protocols were conducted to explain the abstraction of the statistical analyses. In particular, the protocols of those children who answered only structurally (26%) or only semantically (21%) were examined for effects of item (configuration) order or presentation mode order. No such effects could be seen. Nor were there patterns according to grade level or reading level (using both instructional level of the child's reading group and the range of word lists from which the Knownness Rating was drawn as measures of reading level).

An additional use of the individual protocols was to determine the kinds of responses that children made, including their use of metalanguage. Confirming evidence was produced that these children continued to use metalanguage inaccurately even in sixth grade, particularly in reference to structural aspects of words. Two strategies in dealing with word meanings became evident by fourth grade. Children would project a meaning onto an unknown word or would create a hypothetical context for an unknown word. In using these strategies in dealing with unknown words semantically, fourth and sixth grade children used metacognitive language, or language about their thought processes. They used metalinguistic terms less frequently when giving semantic responses than when giving structural responses.

Discussion

In view of previous research it is interesting that the effect of grade level was non-significant. The absence of a main effect for grade when we know children learn more words and word meanings as they grow older underlines the issue of whether there is a single semantic concept corresponding to "word" that would be manifested as a straight developmental trend or whether the notion of word for children depends upon the mode of presentation or other features such as knownness, as indicated by these results. In controlling for word knownness, this study fits with other current research and theory which throws doubt on learner characteristics previously believed to increase (as a main effect) with age (Richman, Nida, & Pittman, 1976; Brown, 1975). The study needs to be replicated and also needs to include more sensitive controls over word knownness, extending to both the oral and written presentation modes.

The descriptive analysis indicated that sixth grade children (and, to some extent, fourth grade children) were able to control strategies of projecting meaning onto unknown words and of creating hypothetical contexts in which unknown words as well as known words would "make sense," using the child's intention as the measure. Examinations of the interactions for written and oral presentation modes with grade level indicate that the means come together at sixth grade. Replication of the study should include an extension upward into adolescence to see if main effects for grade level should appear from sixth grade upward. Indeed it might be predicted that the results for written and oral modes would be the same for all configurations after sixth grade, as they are for K-KK in first through sixth grades. It can be argued that with the known words (K-KK) the child does not have to invoke any meaning creation strategies, whereas, in all three of the other configurations (K-UU, U-KK, U-UU), s/he must.

The descriptive data raise the question of whether children who are reading in school tasks would not benefit from greater understanding of and use of metalanguage. Given the state of research in reading methodology, we have little evidence now. Linnea Ehri's contention (1976) that it is unnecessary to teach metalanguage as a *pre*requisite to reading seems reasonable; children in this study who had learned to read successfully were learning about the referents of metalanguage during and not just prior to their contacts with written language. The only immediate implication that could be drawn from this study for teaching is that teachers should try to learn what their students, individually, understand about and mean by the metalinguistic terms used in school tasks.

References

Anglin, J. M. *The growth of word meanings.* Cambridge, Mass.: MIT Press, 1970.

Bates, E. *Language and context: The acquisition of pragmatics.* New York: Academic Press, 1976.

Brown, A. L. The development of memory: Knowing, knowing about knowing, and knowing how to know. In H. W. Reese (Ed.), *Advances in child development and behavior, Vol. 10.* New York: Academic Press, 1975.

Clark, H. H. The language-as-a-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior,* 1973, *12,* 335-359.

Ehri, L. C. Do words really interfere in naming pictures? *Child Development,* 1976, *47,* 502-507.

Ehri, L. C. Word consciousness in readers and prereaders. *Journal of Educational Psychology,* 1975, *67,* 204-212.

Ehri, L. C. Word learning in beginning readers and prereaders. Paper presented at the American Educational Research Association, New York City, April 1977.

Garvey, C. The contingent query: A dependent act in conversation. In M. Lewis & L. A. Rosenblum (Eds.), *Interaction, conversation, and the development of language*. New York: John Wiley & Sons, 1977.

Gibson, E. J. Perceptual learning and the theory of word perception. *Cognitive Psychology*, 1971, *2*, 351-368.

Golinkoff, R. M., & Rosinski, R. R. Decoding, semantic processing, and reading comprehension skill. *Child Development*, 1967, *47*, 252-258.

Harris, A. J., & Jacobson, M. D. Basic vocabulary for beginning reading. *Reading Teacher*, 1973, *26*, 392-395.

Holden, M. H., & MacGinitie, W. H. Children's conceptions of word boundaries in speech and print. *Journal of Educational Psychology*, 1971, *63*, 551-557.

Karpova, S. N. Osoznanie slovesnogo sostava rechi rebenkom doshkol'nogo vozrasta (The preschoolers' realization of the lexical structure of speech). *Voprosy Psikhologii*, 1955, *4*, 43-55. Abstracted in F. Smith & G. A. Miller (Eds.), *The genesis of language*. Cambridge, Mass.: MIT Press, 1966.

Mason, J. M. *Questioning the notion of independent processing stages in reading*. Technical Report No. 8. Urbana, Ill.: Center for the Study of Reading, 1976.

Mason, J. M. The roles or orthographic, phonological, and word frequency variables on word-nonword decisions. *American Educational Research Journal*, 1976, *13*, 199-206.

Mickish, V. Children's perceptions of written word boundaries. *Journal of Reading Behavior*, 1974, *6*, 19-22.

Naron, N. K. Developmental changes in word attribute utilization for organization and retrieval in free recall. *Journal of Experimental Child Psychology*, in press.

Papandropoulou, I., & Sinclair, H. What is a word? Experimental study of children's ideas on grammar. *Human Development*, 1974, *17*, 241-258.

Richman, C. L., Nida, S., & Pittman, K. Effects of meaningfulness on child free-recall learning. *Developmental Psychology*, 1976, *12*, 460-465.

Rosinski, R. R., Golinkoff, R. M., & Kukish, K. S. Automatic semantic processing in a picture-word interference task. *Child Development*, 1975, *46*, 247-253.

Sheldon basic reading series. Boston: Allyn & Bacon, 1968.

Spache, G. *Good reading for poor readers*. Champaign, Ill.: Garrard, 1968.

The American heritage dictionary of the English language. Boston: Houghton Mifflin Company, 1975.

Thorndike, E. L., & Lorge, I. *The teacher's word book of 30,000 words*. New York: Teacher's College Press, 1944.

Tovey, D. R. Children's perceptions of reading. *Reading Teacher*, 1976, *29*, 536-540.

Wickens, D. D. Characteristics of word encoding. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, D. C.: H. V. Winston & Sons, 1972.

Wickens, D. D. Encoding characteristics of words: An empirical approach to meaning. *Psychological Review*, 1970, *77*, 1-15.

RICHARD L. ALLINGTON
State University of New York
State University of New York at Albany

MICHAEL STRANGE
University of Texas

Word prediction of good and poor readers

The role of semantic and syntactic cues in fluent reading has been the basis for much recent controversy. While some have suggested that effective use of the contextual information provided by these cues is the determining factor in differentiating good and poor readers (Goodman & Goodman, 1977; Smith, 1976) others have argued that virtually all readers, regardless of achievement, employ semantic and syntactic cues and other factors must account for achievement differences (Weber, 1971; Kolers, 1975; Allington & Strange, 1977; Allington, 1978). In each of these latter studies the use of visual information, or an interaction in the use of visual and contextual information seemed to differentiate good and poor readers. Weber (1971), for instance, noted that the majority of errors by all readers conformed to preceding contextual constraints but that good readers seemed to produce errors which more closely approximated visual characteristics of the target word. Similarly, Kolers (1975) found good readers' recognition memory for visual features of sentences exceeded that of poor readers. Allington and Strange (1977) found poor readers ignored visual anomalies in text and gave instead a response which fit syntactic and semantic constraints more often than did good readers who seemed more constrained by the visual information. Finally, Allington (1978) demonstrated that poor readers' recognition accuracy suffered more than good readers' when syntactic information was eliminated.

However, use of context does not seem to be an either/or situation. Mason (1977) has demonstrated the interdependence of various types of processing while reading, a result which supports recent interactive processing models of reading (Rumelhart, 1975; Mosenthal, Walmsley, Allington, 1978). A simplified instructional strategy for inducing such interactive processing has been recently proposed by Dahl and Samuels (1977) and is called 'hypothesis-test' training. Here readers are taught to utilize both semantic-syntactic and grapho-phonic information integratively. Instruction of this type produced a higher level of reading achievement than other more traditional methods (Samuels, Archwamety & Dahl, 1974).

However, we still know little about how readers come to develop a sensitive strategy for employing these information sources. Goodman (1965) and Biemiller (1970) have demonstrated that use of contextual information develops with reading achievement and Pearson and Studt (1975) have demonstrated the positive effects of contextual richness and word frequency upon the word prediction behaviors of readers. There is a particular need to clarify the utilization of semantic-syntactic cues by poor readers. The question, then, for the present study is whether good and poor readers of the same age level differ in performance on a task which requires the integration of semantic-syntactic and grapho-phonic information. Additionally, the performance of older poor readers will be compared to that of younger good readers. These comparisons should also provide an opportunity to test a recent proposal that poor readers have no specific skills deficit but rather perform very much like younger good readers (Guthrie, 1973).

Method

Subjects

The students were drawn from three cooperating schools. In each school all second and fourth grade children were screened for reading ability on the *Peabody Individual Achievement Test.* Students scoring at or above grade level were considered good readers while subjects scoring one year or more below grade level were considered poor readers. From the pools of students available 15 fourth grade good readers (*M* reading grade = 5.7) and 15 fourth grade poor readers (*M* reading grade = 3.3) were randomly selected as were 15 second grade good readers (*M* reading grade = 3.3). Students from three schools were selected in an attempt to minimize specific instructional program effects (Barr, 1977).

Materials

The experimental materials were those employed by Pearson and Studt (1975) and are described in detail there. Briefly, the materials consisted of 36 sentences each with one word deleted and providing three levels of context: rich, moderate, and poor. Two words, of high and low frequency, were designated as target items for each sentence as indicated on the following page.

Word Pairs	Context Level	Sentences
HF Stop	Poor	We decided to for awhile
	Moderate	The men were ordered to
LF Halt	Rich	You had better decide to for that red light

Procedure

Students were tested individually in small rooms adjacent to their classrooms. The experimenter provided a sample sentence explaining that students were to read the sentence and try to think of a word that would make sense in the blank. If the word provided was not the target word then the first letter of the target word would be expressed and they were to read the sentence again and try to think of a word which made sense in the sentence and began with that letter. If an incorrect response followed an additional letter was exposed until either the correct response was elicited or all letters of the word were exposed.

Students were given six sentences, two at each level of contextual richness. At each level the students were asked to provide a high frequency target word for one sentence and a low-frequency target word for the other.

Results and Discussion

A repeated-measures analysis of variance was used to analyze the data. Because there were differing numbers of letters across synonym pairs the proportion of the total word necessary to achieve recognition, rather than the number of letters, was the basic unit of analysis. This was the same unit of analysis used by Pearson and Studt (1975). Cell means and standard deviations are reported in Table I.

Table 1

Cell Means and Standard Deviations: Percentage of Words Needed for Identification

	High Frequency						Low Frequency					
	Rich		Moderate		Poor		Rich		Moderate		Poor	
Students	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
4G	7.2	(1.7)	39.8	(3.2)	49.4	(3.3)	59.9	(2.9)	65.8	(2.8)	80.4	(2.4)
4P	31.2	(3.4)	55.3	(3.9)	69.4	(2.8)	74.9	(2.6)	85.8	(1.7)	95.2	(1.2)
2G	31.7	(3.8)	49.6	(3.5)	69.3	(2.8)	71.0	(3.3)	69.2	(3.3)	88.9	(2.3)

There were significant main effects for each of the factors under consideration. The 4th grade good readers needed 50.45% of the word to achieve recognition, 4th grade poor readers needed 68.66% and 2nd grade subjects needed 63.31%. These differences were significant $F(2,57) = 11.70$, $p < .001$. The mean for high frequency words was 44.79%; for low frequency words 76.82%. This difference was significant $F(31, 57) = 105.18$, $p < .001$. Unlike the Pearson and Studt study, there was no interaction between frequency and group. This would indicate that the effect for frequency was equally distributed over all groups.

There was a significant effect for context, $F(2,114) = 24.53$, $p < .001$. The mean proportion of word necessary to achieve indentification for rich context sentences was 46.0%; for moderate context, 60.9%, for low context, 75.5%. The effect of context differs across the word frequency levels as indicated by the significant interaction between these variables, $F(2, 114) = 4.84$; $p < .01$.

These results, with the exception of the lack of a group x frequency interaction, are similar to the Pearson and Studt (1975) results and their discussion is equally relevant for this study. The fact that 4th grade good readers were able to identify the target words with less graphic information lends credence to their conclusion that the ability to use context is a function of reading proficiency. The context x frequency interaction supports their conclusion concerning response availability. A richer context was more helpful when the target word was a high frequency word. When the target was a low frequency word much more graphic information was needed to achieve recognition. We also found many instances of students supplying the high frequency synonym for low frequency target words.

The second purpose of this study was to determine if good and poor readers employed different strategies to identify unknown words. In order to answer this question the incorrect responses were analyzed to infer which cue systems were being used to arrive at a response. In the absence of graphic cues, the most frequent choice of all groups was to give a response that was semantically and syntactically appropriate ($M = 81.0$). Second grade students gave such a response 75.7% of the time, 4th grade poor readers gave such responses 75.9% of the time and 4th grade good readers gave such responses 91.5% of the time.

Once graphic cues became available the response pattern changed. The most popular strategy was to give no response (*M* = 51.0) with this occurring in second grade students 49.8% of the time, in 4th grade poor readers 51.7% of the time, and 4th grade good readers, 51.6%. The second most popular strategy was to give a response that was graphically, semantically and syntactically appropriate (*M* = 27.3%). Second grade students gave such responses 26.9% of the time, 4th grade poor readers 23.7% and 4th grade good readers 31.2%. In each of these analyses the older poor readers and the younger readers performed similarly, responding less frequently to semantic and syntactic constraints than the older good readers.

Conclusions

The response analysis indicates that all students were able to utilize graphic, syntactic and semantic cues. Good and poor readers do, however, seem to differ on the integration of these cue systems. This would support a notion of the reading process that includes an increase in the ability to integrate the cue system as a function of an increase in reading fluency (Pearson & Studt, 1975). In this respect this study also supports Guthrie's (1973) proposal that older poor readers perform like younger good readers. In the absence of visual information (Trial 1) students in all groups were likely to supply meaningful responses, that is responses that were syntactically and semantically appropriate. Once graphic information was introduced (Trials 2-4) each group seemed to be affected in somewhat the same fashion. Close to 50% of the students in each group chose a no response strategy. This would seem to further support Pearson and Studt's (1975) conclusion concerning response availability. It seems that the students were able to achieve an acceptable meaning for the target word but the fact that they lacked a word that fit all the requirements (graphic, syntactic and semantic) inhibited their ability to respond. This conclusion is confounded since the students knew their first response was in some way inappropriate. If a response was given, most incorporated the graphic information. In fact, it would seem that graphic acceptability became the most salient cue in selecting a response with the better readers who were more likely to produce a response that agreed not only with this information, but also with each of the other available cues. However, as noted earlier, the introduction of graphic constraints inhibited responses for the less skilled readers.

In summary, then, the older good and poor readers did differ on their ability with the experimental task but there seem to be no differences in the performances of the older poor readers and younger good readers. All groups were affected by contextual richness and word frequency but in some cases to different degrees. Finally, more skilled readers seemed to be able to use graphic information in conjunction with contextual constraints more efficiently and effectively than the less skilled readers.

References

Allington, R. Effects of contextual constraints upon rate and accuracy. *Perceptual and Motor Skills*, 1978, *46*, 1318.

Allington, R. & Strange, M. Effects of grapheme-substitutions in connected text upon reading behaviors. *Visible Language*, 1977, *11*, 285-297.

Barr, R. The influence of instructional conditions upon word recognition errors. *Reading Research Quarterly*, 1972, *1*, 509-529.

Biemiller, A. The development and use of graphic and contextual information as children learn to read. *Reading Research Quarterly*, 1970, *6*, 75-96.

Dahl, P. & Samuels, S. J. Teaching children to read using hypothesis/test strategies. *Reading Teacher*, 1977, *30*, 603-606.

Goodman, K. & Goodman, Y. Learning about psycholinguistic processes by analyzing oral reading errors. *Harvard Educational Review*, 1977, *47*, 317-333.

Goodman, K. A linguistic study of cues and miscues in reading. *Elementary English*, 1965, *42*, 639-643.

Guthrie, J. Models of reading ability and disability. *Journal of Educational Psychology*, 1973, *63*, 9-18.

Kolers, P. Pattern-analyzing disability in poor readers. *Developmental Psychology*, 1975, *11*, 282-290.

Mason, J. Questioning the notion of independent processing stages in reading. *Journal of Educational Psychology*, 1977, *69*, 288-297.

Mosenthal, P. M., Walmsley, S. A. & Allington, R. L. Word recognition reconsidered: Toward a multi-context model. Paper presented at the International Reading Association, Houston, 1978.

Pearson, P. D. & Studt, A. Effects of word frequency and contextual richness on children's word identification abilities. *Journal of Educational Psychology*, 1975, *67*, 89-95.

Rumelhart, D. Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and Performance*, VI. London, Academic Press, 1975.

Samuels, S. J. & Archwamety, T. & Dahl, P. Effect of hypothesis-test teaching on reading skill. *Journal of Educational Psychology*, 1974, 66, 833-844.

Smith, F. The role of prediction in reading. *Language Arts*, 1975, *52* 305-311.

Weber, R. A linguistic analysis of first-grade reading errors. *Reading Research Quarterly*, 1970, 5, 427-451.

GREGORY PEARSON
Cornell University

The acquisition of knowledge from text

A progress report on prose research at the Cornell University reading research group (*)

Research on the acquisition of information from texts through the process of reading is of obvious theoretical and practical importance. If we could specify how the structure of the information in a text passage affects the structure of the knowledge acquired from that text then we would not only be able to say a great deal about cognitive functioning but we would be able to make a number of statements of great practical value as well. Among other practical statements, we might be able to specify how a text should be organized to enhance the information acquisition process or to characterize questions to appropriately probe the knowledge acquired by the reader.

Clearly the motivation for prose research of this type is very high. This is fortunate for research in this area is frustrated by the sheer enormity of the research domain, not to mention the complexity of the issues involved. Linguists, for example, have taken a long time to get around to solving problems of semantics, and for good reason. The study of meaning, and the acquisition of meaning, is a very difficult task. To make matters worse, there is even a lack of good tools for use in this kind of prose-related research.

In this paper I would like to be able to report that we have solved a lot of these problems in a series of brilliant studies. Unfortunately this is not the case. During the past two years, however, we have made considerable progress in the development of research tools and we have used these in several studies on which I will briefly report later.

An Information Processing Model

To begin, let's conceptualize the knowledge acquisition process as a problem in data communications, as shown in Figure 1. Some individual, an author, searches his base of data about the world and selects from it a set of data items which is to comprise the content of his written message. At this first stage the author must make various pragmatic decisions about what information to include in the content structure based on considerations such as the expected state of knowledge of a potential reader and about the inferential capabilities of that reader. The second state in this process model is the organization of the selected data or semantic content into a coherent text base from which a natural language representation may be generated. At this stage the information must be ordered into a reasonable sequence and various parts of it marked for discourse features such as topicalization and focus. Finally linguistic processes must operate on this text base to produce the actual natural language strings which constitute the written text.

The reader, presented with a text, must reverse this generation process. He must decode or parse and semantically interpret the surface form of the

---

(*) The research reported here was supported through grant number NIE-G-74-0018, Structure and Learning From Prose, from the National Institute of Education, awarded to Dr. George W. McConkie.

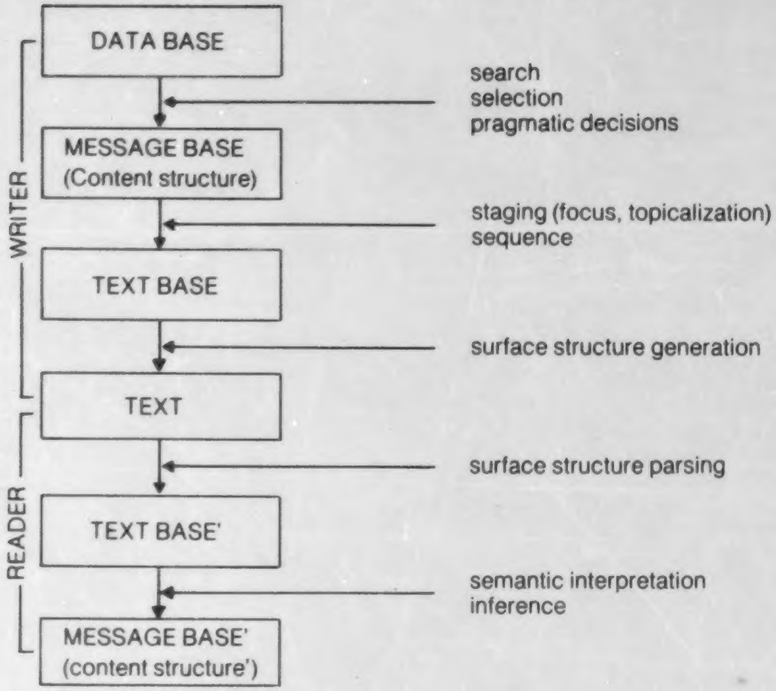


Figure 1. An information processing model of text generation and text decoding.

text's natural language representation and generate a text base of his own. From this he then must apply any relevant inferential processes to arrive at the content structure which he concludes must have been the source of the author's written text.

Note that the model here is not necessarily a psychologically real one but is rather an attempt to identify the kind of processing which semantic information must undergo to conform to something that we readily know. We can only directly observe one of these stages, the text itself. The rest we must infer. But since people can freely generate texts on all kinds of topics in infinite variety, we must postulate a process of selecting from the writer's knowledge what he wants to say, or the *message base*. And since the same set of information can be expressed in many different ways, different styles or with different focus, we must hypothesize a *text base*, containing such detail about how the text will be constructed. Below I will mention a study done in our group which gives evidence for the effect of staging operations and thus for existence of a text base stage.

The Representation of the Content Structure Level

There is an obvious need in prose-related experimental research to describe and characterize the stimulus, the text passage to be read. Also, since in our

72

studies we have used mostly free recall protocols, that is having the subject write down everything he can remember from a passage, we also need ways of characterizing these free recalls, so we can identify just which aspects of this original passage have been retained. Clearly these descriptions must be done in a non-arbitrary manner and in such a way so as to avoid the extreme variability which is evident in the surface structure language of any of these texts. This method of description is one of the research tools which has been mostly missing but sorely needed for research of this type. Consequently we have expended and are continuing to expend a considerable amount of energy on the development of a consistent and reliable method of representing the semantic information of a text at the content structure level, the first level seemingly independent of linguistic processing.

Some time ago we began using a propositional network for representing information at this level. In a network notational system of this type, the content or meaning of a text is divided up into underlying semantic units, primarily states and events, and each unit is represented by a set of nodes and arcs connecting the nodes. Each node contains a semantic concept while the arcs are labelled to represent the pair-wise relationships which exist between the various concepts contained in the nodes. In each case, be it an event or a state, the resultant set of nodes and arcs is called a proposition.

As a very simple example of this kind of representation, consider the following two propositions given as an example in McConkie (1977):

Proposition 1: (:BALL)–ATT-> ('RED) ATT = attribute
(a state)
Proposition 2: (:JOHN)–AGT-> ('HIT)–OBJ-> (:BALL)
(an event) AGT = agent
OBJ = object

As McConkie points out, these two propositions constitute a content structure which could be said to represent the meaning of a number of different test strings such as

John hit the red ball.
The red ball was hit by John.
The ball hit by John was red.
The ball which John hit was red.
It was a red ball which John hit.
The red ball which John hit...
etc.

Thus, the network represents certain semantic relations which can be expressed in many different ways, and ignores many aspects of variability in the form in which that information is expressed.

The propositional notation system which we currently use is a somewhat simplified version of the one described in Frederiksen (1975). Frederiksen's system was chosen over other representational systems for several reasons. First, his system is highly detailed. By incorporating a large number of fine distinctions in his taxonomy of concepts and in the classification of the relations which can co-occur with these various concept classes, his system is potentially capable of representing a broad range of semantic information. Secondly, since a network representation has a natural implementation in a computer as a list structure, Frederiksen's notation is suitable for use as the data structure in a computer-based data handling system. This is an important consideration since one of the things that makes prose research difficult is the great ease with which one can become innundated with data. Some kind of automation of the data handling requirements is really mandatory. In fact, we have developed an interactive lab computer software system, called PROSYS, for facilitating the input of such networks and for operating on them for data handling and analysis functions.

Finally, Frederiksen's system is suitable for hand-coding. I use the term 'hand-coding' here to point out the difference between the way we arrive at the content structure of a text and the approach taken by a number of workers in the field of artificial intelligence. In our case, human scorers accomplish the analysis of a text into a propositional network which represents the content structure of that text. In the artificial intelligence case a computer program parses an input text and renders it into a semantic representation. In a real sense the output of the two different approaches is the same - a semantic representation at something like what I have been calling the content structure level.

If the two approaches output essentially the same thing, then why not employ an automatic parsing approach? The reason is quite simple - as yet such computer text understanding systems are not as good as human scorers at the kind of semantic analysis that is required for psychological research. The problem with automatic parsing is that many natural language text strings present enormous difficulties for machine parsers while human scorers can easily understand and represent their meanings. Consider the following example from Charniak and Wilks (1976). Suppose a restaurant patron makes the following statement: "Waiter, I would like spaghetti with meat sauce and wine." The patron would be surprised to get spaghetti covered with a mixture of meat sauce and wine. If the order had been 'with garlic and butter,' however, a mixture of the two is exactly what would have been expected. Charniak and Wilks point out that this has been considered to be a linguistic problem, one of structural ambiguity of the underlying syntax of the 'with' phrase. The meaning of this phrase, however, is disambiguated on other than syntactic grounds. An inference must be made based on one's knowledge of food. This kind of disambiguation process is simple for people and a scorer would have no difficulty in representing the meaning of this sample sentence. A machine parser, on the other hand, can have problems with ambiguities of this type.

Of course there are problems with hand-coding as well. One of the most significant is the problem of achieving good inter-scorer reliability. Unreliability comes from two sources: lack of agreement on the meaning to be ascribed to a specific piece of text and the use of different representations for the same meaning. We have been working to reduce this latter source of unreliability by developing explicit and detailed descriptions of the elements and relations available in our modified propositional notation system and by producing a set of procedures for using these. Pearson (1977) reports on the reliability of using a propositional network representation in a slightly different case, namely in a situation where the network representation of a stimulus text has been determined and is to be used for the scoring of free recall protocols. The reported results indicate that the degree of reliability vaires with respect to the level of analysis. When considering the amount of information recalled in the protocol, that is, the number of elements in the stimulus text representation found to be present in the recall, the reliability was extremely high (r = .99). When considering the type of information recalled, the reliability varies with the level of representation, that is, with the number of categories (for propositions r = .84, for elements r = .78). It is interesting to note that in this reliability study the largest amount of variability, some 17%, was due to scorer errors or omissions and that only 7% was attributable to notation-related difficulties.

Some Examples of Research

In the information processing model (Fig. 1) discussed above, I have depicted 'staging' as an operation on the message base which identifies the relative prominence given to various segments of the semantic content. There are a number of devices writers use to vary the staging level or degree of prominence given to different information in the text. In one study from our laboratory, Paul Clements (1976) investigated what effect some of these have on the reader, that is, what is the role of staging in the reception of information. In these experiments two groups of students were given pairs of passages which differed with respect to the level (high prominence vs. low prominence) at which specific chunks of information were staged. The hypothesis was that specific pieces of information would be better recalled after reading a passage in which that information is staged high than after reading a passage in which the same information was staged lower. The experimental results strongly supported this prediction and further suggest that staging primarily effects the acquisition of information during reading itself rather than having its effects only later during retrieval. As part of the experimental procedure, Clements had his subjects rate pieces of information in the passage for their relative importance. The results showed that subjects' ratings did not correlate with staging differences. Thus, even though information staged higher was not recognized as being more important to the passage, still the readers were more likely to remember it.

In another study, Lucas (1977) investigated whether the type of test format that readers anticipate influences what they retain from reading a passage. Subjects were asked to read a series of two passages. After the first passage, one group was given a number of test questions and another group was asked to write down everything they could remember - a free recall test. Subjects expected the same type of test on the second passage, but instead both groups were given a free recall test. The experimental results showed that the type of test a group expected did have an effect on what they retained from the passage. The most obvious effect was that subjects who were given a free recall test after the first passage read the second passage an average of 28% slower than those who had been led to anticipate the question test, although they did not recall significantly more propositions. Lucas notes that

> A difference in reading time would seem to indicate a difference in the processing activities of the readers, which should result in differences in what is retained from the passage. While it is clear that characteristics of the passage itself had the greatest influence on what was recalled, the reading strategies did have some effect, as revealed by the results of the analyses on patterns of recall information. The fact that the total number of propositions recalled did not differ significantly between the experimental groups suggest that had the recalls been analyzed using "idea units" or other less fine-grained techniques, the conclusions drawn might well have been "no difference."

Data from this further analysis led Lucas to conclude that anticipation of a free recall test may influence a reader to acquire more central aspects of a text and to be more aware of the manner in which the information was presented in the text.

The final study which I will briefly mention is one by Scott Smith (1977) which investigated memory over varying intervals of time. Subjects were given a passage to read and immediately produce free recalls. After three weeks some of the subjects were unexpectedly tested again and another group was tested after three months. Smith's most significant finding was that

> First, the recalled information most similar to the passage is much more stable over time than recalled information modified from the passage. Second, there is a substantial amount of information present in delayed recalls that was not present in an immediate recall. Thus, the immediate recall must not provide a complete indication of the information which the subject has retained from the passage.

In his analysis, Smith was able to see which concepts and propositions, at the lowest level of analysis, were retained over different retention intervals. However, none of the studies mentioned have made attempts at comparing the structure of the content of the passages read with that of the recall protocol.

Summary and Promise for the Future

What then do we know from all of this? First we conclude that in order to answer many of the detailed prose-related questions, we need to have the kind of rich experimental data which we get from using a detailed propositional network representation of semantic content structure. These are questions like

- -what types of influence does the passage structure have on what readers tend to remember from it?
- -what are the structural characteristics of subjects' cognitive representations of information from the passage, and how do text and task variables influence this?
- -in what ways do passage and cognitive structure influence the retrieval and use of information obtained from passages?

Second, we know we need to further specify procedures for analyzing semantic content structures and third, we need some computer tools to help us pull structural information from the huge data matrices which result from the translation of semantic networks into their equivalent numerical form. These, of course, are primarily methodological 'results'. We believe that the analysis techniques and computer programs which we are developing will be important contributions to this type of research. In addition, I have mentioned some substantive things which have been learned, the most important of which is the work by Paul Clements on staging.

The promise for the future is for more small steps which hopefully will chip away at the big problem. Recently we have been making good progress in the third methodological area I just mentioned - the task of pulling structure out of recall semantic content structures. With good computer tools for this, there is promise of additional experimental results which will hopefully move us closer to answers to the kinds of theoretical and practical questions with which I started.

References

Charniak, E. & Wilks, Y. *Computational semantics.* New York: North-Holland, 1976.

Clements, P. *The effects of staging on recall from prose.* Doctoral dissertation, Cornell University, 1975.

Frederiksen, C. H. Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology, 7,* 1975.

Lucas, P. A. *Anticipation of test format: Some effects on retention.* Paper presented at the annual meeting of the American Educational Research Association, New York, April 1977.

McConkie, G. W. *Studying retention from prose: A content structure approach.* Paper presented at the annual meetings of the American Educational Research Association, New York, April 1977.

Pearson, G. *Representing meaning in text.* Paper presented at the annual meeting of the American Educational Research Association, New York, April 1977.

Smith, H. *Memory over varying intervals of time.* Paper presented at the annual meeting of the American Educational Research Association, New York, April 1977.

CONNIE A. BRIDGE
University of Kentucky

ROBERT J. TIERNEY
University of Arizona

MARY JANE CERA
Kino Learning Center

Inferential operations of children involved in discourse processing

The factors involved in the comprehension and memory of text have recently become the focus of studies in verbal learning. Basic to the study of text comprehension has been the development of detailed models of discourse structure, which are seen as a representation of the knowledge structure of the writer or speaker producing the text. The ability to analyze a text into specifically defined and ordered units has provided networks against which the subject's memory for text may be compared. This comparison of the subject's knowledge structure acquired by reading the text to the knowledge structure underlying the generation of the text itself has provided insight into the characteristics of discourse as well as the interaction between textual characteristics and discourse processing operations.

One aspect of discourse processing that becomes obvious when the network of the subject's recall is compared to the network of the original text is that when subjects read and recall text, they acquire some information which was not explicitly stated in the text itself. This generated or inferred information can be investigated by comparing the propositions in the subject's recall which were not explicitly stated in the text to those propositions as actually stated in the text.

An examination of the inferences and inferential processes involved in generating inferred propositions from stated ones provides supportive evidence for constructivist theories of language processing. The constructivists contend that the comprehension-memory system does not process each input sentence in its entirety; but, rather, selectively processes the input, using information selected from the input sentence and stored knowledge about the world to generate a semantic interpretation based upon the input data in the text (Frederiksen, 1975b).

The definition of inference used in this study was relatively unrestricted in order that it might reflect the pervasive role of inference in language processing. As stated by Frederiksen (1977b), "Inference occurs whenever a person operates on semantic information, i.e., on concepts, propositional structures, or components of propositions, to generate new semantic information, i.e., new concepts or propositional structures. Any semantic knowledge which is so generated is *inferred*" (p. 7). Thus, any information generated by the reader that was not explicitly stated in the text was classified as an inference.

The purposes of this study were to find out how much explicitly stated information subjects remembered from reading a passage, how much inferred information they generated during that process, how many total units of information were recalled, and the relative proportion of inferred information to total information in their recalls. Furthermore, the study was designed to see if there was a difference between the groups of good and poor readers in regard to these factors. Finally, the inferred information was analyzed by type of inference and class of inferential operation employed in generating the inference to discover how many inferences of each class and type were generated by each subject individually and by each group of good and poor readers and further to see if there was a difference in the proportionate use of the eight classes of inferential operations employed by the two groups.

Method

Subjects

The subjects for this investigation were 36 third grade children from the Sunnyside School District in Tucson, Arizona. Eighteen of the students were classified as good readers, while the other eighteen were termed poor readers. The assignment of students to these categories involved two criteria: teacher judgment and Total Reading Scores on the *Stanford Reading Achievement Test.*

Procedure

The students were asked to read a 153 word non-narrative, informational passage. They were told that after the reading they would be asked to tell everything they could remember about the selection. Between the reading and the retelling of the target passage, they were asked to read orally a short narrative selection which was to serve as a buffer to short term memory.

The entire session, in which the child read the target passage and the buffer passage orally and retold what he remembered in a free recall followed by a probed recall, was tape recorded. The semantic and logical network of each recall was then compared item-by-item to the message base of the original passage. Using the semantic and logical network of the passage as a scoring key, every item that appeared in the semantic and logical network of the recall was given one point, with each concept and each relation scored as a separate item. The score obtained by summing the above items represented the amount of explicit or reproduced information in the recall.

Analysis

The items of information appearing in the text base of the recall which were not in the message base of the passage could then be counted and classified according to Frederiksen's (1977a) Taxonomy of Text-Based Inferences. Each inference was first analyzed by comparing the inferred proposition to the original proposition. This comparison permitted its classification according to one of the twenty-six inference types and consequently into one of the eight classes of inferential operations.

A check of interrater agreement on the scoring of the recalls in terms of amount of explicit and inferred information revealed a relatively high percentage of agreement between scorers (.903). However, the system of classifying inferences appears to be capable of yielding data of only moderate stability at this point with the raters agreeing on .824 of the inferences classified.

Results

1. Good readers recalled significantly more units of explicit information, $t(34) = 5.84, p < .01$, generated more units of inferred information, $t(34) = 2.87, p < .01$, and had longer total recalls, $t(34) = 4.74, p < .01$. (See Table 1)

Table 1

Mean number of explicit, inferred, and total
items of information in the recalls of good and poor readers

Types of Information	M	SD
Explicit Items		
Good	91.89	25.46
Poor	48.17	19.01
Inferred Items		
Good	85.61	51.17
Poor	44.78	31.87
Total Items		
Good	177.50	60.36
Poor	92.94	45.61

2. Within each group and for each type of information (explicit, inferred, and total), there was a wide range between the subject with the highest number of units recalled and the one with the lowest.

3. There was considerable overlap in the scores, with some of the poor readers recalling more units of information than some of the good readers.

4. In both groups there were individual subjects who had high explicit scores but low inferred scores. The reverse situation also occurred, but less frequently.

5. The individual who had the highest explicit recall score remembered only half of the total possible concepts and relations in the passage.

6. There was no significant difference in the proportion of inferred information to total information in the recalls of good readers compared to poor readers.

7. Good readers made significantly more causal and conditional inferences, both of which involved dependency operations, t (34) = 3.66, $p < .01$, (see Table 2).

8. Poor readers made significantly more superordinate inferences in which a more general concept was substituted for a specific one. These overgeneralizations involved macrostructure operations, t (34) = −3.45, $p < .01$.

9. There was no significant difference in the proportion of inferences in the other six classes of operations included in the Taxonomy of Text-Based Inferences.

Conclusions and Discussion

Based upon the findings regarding amount and type of explicit, inferred, and total information in the recalls, it was possible to conclude that:

1. Both good and poor readers engage in some constructive or inferential processing during reading and recall, with nearly half of the information in the retelling being inferred.

These data support the constructivist view of comprehension which holds that the reader selectively processes the input data, using information from the text in association with his/her own background of knowledge to construct a meaningful interpretation of the material. Because of this interaction of previous knowledge with the new information, it was not surprising that as the recall of explicit units increased, so did the generation of inferred units.

As Anderson, Reynolds, Schallert, & Goetz (1976) pointed out, in terms of schema theory, the main determinant of the knowledge that an individual is able to acquire from reading is the knowledge that he/she brings to the task. If a poor or inexperienced reader does not possess the relevant schema or does not know how to bring this schema to bear upon the task, he/she will be unable to comprehend the material.

Table 2

Mean proportion of inferences in each class
of inferential operations in the recalls of good and poor readers

Class of Inferential Operation	Mean Proportion
Class I Lexical	
Good	.0200
Poor	.0194
Class II Identifying	
Good	.2161
Poor	.2088
Class III Frame	
Good	.1068
Poor	.1080
Class IV Event	
Good	.2011
Poor	.1961
Class V Macrostructure	
Good	.0436
Poor	.1871
Class VI Algebraic	
Good	.1918
Poor	.1732
Class VII Dependency	
Good	.1815
Poor	.0746
Class VIII Truth-Value	
Good	.0391
Poor	.0329

2. Comprehension is abstractive as well as constructive, with the readers choosing to retell what they consider to be the most relevant units from the text.

Even the subject who had the highest explicit recall score remembered only half of the total possible concepts and relations in the passage. This provides further evidence for the selective nature of comprehension and recall. In terms of the constructivist theory of comprehension, it seems that the reader selects from the input data those propositions which he/she considers relevant in terms of his/her own schematic expectations and background of knowledge.

3. The comprehension process appears to be essentially the same for both good and poor readers. Both good and poor readers engage in a substantial amount of inferential processing, differing primarily in terms of the good readers' superior ability to deal with interpropositional relationships.

The finding of this study that good readers made significantly more causal and conditional inferences provides evidence of their superior ability to use schematic expectations of relationships to infer unspecified information. Like the truly-fluent adult readers in Marshall's (1977) study, the good third grade readers were able to infer unspecified relational information. Marshall concluded that this awareness of the importance of interpropositional relationships and the use of these relationships to organize their recalls is indicative of the fluent readers' ability to recreate the author's schema.

In summary, it is obvious that an extensive amount of inferential processing occurs during the processing of even "simple" stories, such as the one used in this study. Furthermore, all readers can and do engage in inferential processing during reading and recall, as inferencing is a requisite and integral part of all language comprehension. Good readers generate more inferred information but also recall more explicitly stated information; thus, the proportion of inferred information to total information is almost the same in the recalls of good and poor readers. Good readers generate more causal and conditional inferences between propositions which adds to the coherency and cohesiveness of their recalls, while poor readers tend to substitute more general concepts for the specific terms used in the passage, resulting in recalls which lack the accuracy and specificity of the recalls of good readers.

References

Anderson, R. C., Reynolds, R. E., Schallert, D. I. & Goetz, E. T. *Frameworks for comprehending discourse*, Technical Report No. 12. Urbana, Illinois: University of Illinois, Laboratory for Cognitive Studies in Education, July 1976.

Frederiksen, C. H. Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology*, 1975b, 7, 371-458.

Frederiksen, C. H. Effects of context-induced processing operations on semantic information acquired from discourse. *Cognitive Psychology*, 1975a, 7, 139-166.

Frederiksen, C. H. Discourse comprehension and early reading. In L. Resnick & P. Weaver (Eds.), *Theory and practice of early reading*. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1977a.

Frederiksen, C. H. *Inference and the structure of children's discourse*. Paper for the Symposium of the Development of Discourse Processing Skills, Society for Research in Child Development, New Orleans, 1977b.

Marshall, N. *Comprehension of connected discourse: A study into the relationships between the structure of text and information recalled*. Unpublished manuscript, University of Virginia, 1977.

JILL EDWARDS OLSHAVSKY
Indiana University

Comprehension profiles of good and poor readers across materials of increasing difficulty

The ability of good and poor readers to adjust their strategies to accommodate increasingly difficult material was investigated in this exploratory study. A strategy was defined as a purposeful means of comprehending an author's message. Information processing research by Goodman (1969), Smith (1973), and Newell and Simon (1972) lead to a view of the reader as an active information processor whose interest in comprehending the author's message causes him to apply reading strategies. The ease with which the reader comprehends may be understood in terms of the adequacy of his strategies and the problems set forth in the text. It was expected that good readers given material of varying difficulty, would use strategies more frequently with the most difficult material. Research on the differences between good and poor comprehenders (Smith, 1967; Golinkoff, 1975-1976) indicates good readers are more adaptive and flexible in their patterns of reading than poor readers. Olshavsky (1976-1977) found that good readers apply strategies more frequently than poor readers. Therefore, the hypotheses for the study predicted that only the good readers would adjust their use of strategies according to the difficulty of the material.

H_1: The frequency of strategy usage by good readers will increase as the material becomes more difficult.

H_2: There will be no difference in the frequency of strategy usage across the material by poor readers.

Method

Design

The study utilized a 2x4 analysis of variance with repeated mesures. The first variable was reading proficiency with two levels, good and poor; the second variable was material difficulty with four levels.

Subjects

Twelve eleventh grade students, six good readers and six poor readers were randomly selected from two English classes at a high school in a small midwestern city. Reading ability was determined by scores on the comprehension subtests of the *Stanford Diagnostic Reading Test*, Level III. Good readers were defined as those who scored at stanines 8-9; poor readers scored at stanines 3-4.

Materials

Each student read the same material, excerpts from four short stories. The stories were selected according to readability (levels 7-8, 9-10, 11-12, and 13-15 according to the Dale Chall 1948 readability formula), length (approximately 500 words), and writing style (abstract as rated by five out of six raters).

Data Collection

Each student was instructed to read silently and pause after each independent clause (marked by a red dot) to talk about what happened in the story and about what s/he was doing and thinking as s/he read. Next s/he was given a practice session to familiarize him/her with the procedure and tape recording. The session continued until the researcher decided the student was talking freely. Finally, the student was given each story in rotated order and was told to read each and follow the reading-varbalizing procedure s/he had practiced. The student was not prompted by the researcher as s/he read.

Data Analysis

Matching protocols and clauses. The students' recorded verbalizations were transcribed, and the transcriptions were matched with the corresponding clause in the story. An example follows:

Clause:	He makes no sound at all
Protocol:	The boy's unconscious

Identifying and categorizing strategies. Each protocol, the transcribed record of a student's verbalizations about his/her ongoing behavior, was analyzed to infer reading strategies. Newell and Simon state, ". . . if the subject employs definite processes, . . . there may occur enough repetitions of essentially the same situation to allow us to induce what the processes are and to have some faith in the reality" (1972, p. 191). The process of categorizing students' protocols involved three steps. First, the researcher compared each response with the text and recorded a description of the student's protocol. Second, the researcher decided whether the protocol fit the definition of a strategy. Third, each strategy was given a descriptive name if it occurred at least four times. The protocol presented in the example above was categorized as a strategy because it represented a way of dealing with the author's message. The strategy was termed *inference* because the student added an interpretation of the clause. A second scorer classified the protocols of two randomly selected students with 95% agreement.

Testing the hypotheses. Analysis of variance with repeated measures was used to test the frequency of strategy usage across the four stories.

Determing accuracy of paraphrases. The protocols which were paraphrases of the clause were scored as accurate if they preserved the author's meaning or inaccurate if they distorted the author's message.

Results

Reading Strategies

The eleven reading strategies which were identified are listed in order of frequency and defined: (1) *inference* (125 occurrences) is an addition of an interpretation or suggestion to the information in the story; (2) *hypothesis* (63 occurrences) is a prediction of the meaning; (3) *stated failure to understand a word* (54 occurrences) is an identification of the problem of inability to understand a word; (4) *synonym substitution (48 occurrences)* is use of a synonym in place of a word in the text during the protocol; (5) *personal identification* (42 occurrences) is drawing personal associations in comparison to information in the text; (6) *re-reading* (22 occurrences) is a statement that the clause was read twice; (7) *addition of information* (22 occurrences) is addition of factual information to the information contained in the clause; (8) *skip a word* (14 occurrences) is a statement that a word was skipped; (9) *stated failure to understand a clause* (15 occurrences) is an identification of inability to comprehend the meaning of a clause; (10) *picture story* (5 occurrences) is a statement of a visual image of the story; (11) *conclusion* (4 occurrences) is a statement which summarizes the story.

Tests of Hypotheses

Analysis of variance based on all strategies found no significant difference for the first variable, reading proficiency, $F(1,10) = 2.05, p > .05$. The second variable, difficulty of material was significant, $f(3,30) = 3.71, p < .05$. The interaction effect was not significant, $F(3,30) = .94, p > .05$. The results of this analysis fail to support the first hypothesis but do support the second hypothesis. The total number of strategies used by both good and poor readers decreased (153, 147, 81, 82 respectively) as the stories increased in difficulty.

Comprehension of the Stories

As a check on the effect of the material difficulty variable, an analysis of variance was performed on the overall number of accurate and inaccurate paraphrases by good and poor readers for each story. The test revealed that as the difficulty of the stories increased, the number (23, 18, 35, 62) of inaccurate paraphrases increased, $F(3,30) = 5.5, p < .001$, and the number (325, 185, 133, 120) of accurate paraphrases decreased, $F(3,30) = 28, p < .001$. In neither analysis was the interaction effect significant.

Discussion

The finding that both good and poor readers used the eleven strategies less frequently as the stories became more difficult may indicate they judged their material early in their reading and 'gave up' trying to understand the more difficult selections. In previous research, it was found that good readers used strategies more frequently than poor readers. For this reason, it is assumed that the difficulty of the material rather than the procedure caused these subjects to have similar profiles of strategy usage. The protocol technique, though possibly causing some interference with the reading process, has been used widely in cognitive psychology research because it reveals process and has provided valuable information about reading. Further research should investigate whether the process of applying strategies is a necessary part of comprehension. If strategies are essential, teaching readers to apply them would be one means of helping them cope with the difficult material they encounter in high school and college reading.

References

Dale, E. & Chall, J. S. A formula for predicting readability. *Educational Research Bulletin*, 1948, 27, 11-20, 28.

Fareed, A. Interpretive responses in reading history and biology: An exploratory study. *Reading Research Quarterly*, 1971, *6*, 493-532.

Golinkoff, R. M. A comparison of reading comprehension processes in good and poor comprehenders. *Reading Research Quarterly*, 1975-1976, *11*, 623-659.

Goodman, K. S. Reading: A psycholinguistic guessing game. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading*. Newark, Del.: International Reading Association, 1969.

Newell, A. & Simon, H. A. *Human problem solving*. Englewood Cliffs, New Jersey: Prentice-Hall, 1972.

Olshavsky, J. E. Reading as problem solving: An investigation of strategies. *Reading Research Quarterly*, 1976-1977, *12*, 654-674.

Smith, F. *Psycholinguistics and reading*. New York: Holt, Rinehart, and Winston, 1973.

Smith, H. K. The responses of good and poor readers when asked to read for different purposes. *Reading Research Quarterly*, 1967, *3*, 53-84.

KAREN M. HANSON
Normandale Community College
Bloomington, Minnesota

ROBERT SCHREINER
THOMAS J. HUMMEL
University of Minnesota

The relationship between reading ability and semantic verification tasks

Recently, writers in psycholinguistics and computer simulation of human thought processes have addressed themselves to the processes involved in understanding written discourse. These efforts are particularly apparent in the area of semantic information processing, one means by which readers gain information from written discourse. The present study was concerned with the factors outlined by semantic processing research which may affect the reading comprehension abilities of good and poor adult readers.

The terminology for defining semantic processing in this study was derived from proposals of linguists that the universal basis for semantic memory is not the surface structure of a sentence, despite its effect on the ease or difficulty of comprehending discourse, but rather the deep structure or conceptual organization of the discourse. Fillmore (1968) proposed case grammar as a means of describing deep structure. In this framework the verb serves as the focus of a thought unit and other concepts within written discourse are related to the verb.

Paralleling the work of Fillmore, other researchers attempted to create computer simulated programs that would meaningfully process and respond to human discourse. Frijda (1972) and Minsky (1968) outlined the development of artificial intelligence programming, proposing that schemata for memory structure are similar to the processing strategies of the computer when it responds "intelligently" to human discourse.

The research of linguists and computer scientists converged in Quillian's (1968) proposal for a hierarchically structured semantic network. This network posited by Quillian served as the basis for describing semantic memory processes for human and artificial intelligence. The hierarchical nature of such a network was experimentally verified by Collins and Quillian (1969) when they demonstrated that the time required to recognize a superordinate of a given noun was a function of the hierarchical "distance" between the stimulus word and the superordinate. Furthermore, they found that attributes specifically related to a target noun were more rapidly accessed than were attributes related to nouns superordinate to the target noun in a semantic hierarchy. Additional support for a schematic relationship was developed by: Rumelhart, Lindsay and Norman (1972), Meyer (1970), and Collins and Quillian (1970).

Within this same period, a considerable body of experimental evidence was gathered in support of these hierarchical, constructive theories of human memory. Several studies supported the notion that memory consists of organized relations (Battig & Montague, 1969; Grober & Loftus, 1974; Loftus, 1973; Meyer, 1970; Rosch, 1973, 1975). Other studies demonstrated that

---

* The research reported here is based on the first author's doctoral thesis completed at the University of Minnesota, Minneapolis, Minnesota, 1977.

87

concepts are the basis of memory (Bransford, Barclay & Franks, 1972; Bransford & Franks, 1971).

Freedman and Loftus (1971), Loftus and Grober (1973), and Schaeffer, Lewis and Van de Car (1971) determined the developmental nature of semantic schemata. This finding was important to the present study because it suggests the possibility that one of the differences between good and poor adult readers may result from developmental differences in their semantic encoding and processing strategies.

The goal of the present study was to integrate two paradigms in a single experiment: to combine some traditional reading measures with some novel semantic processing measures in an effort to determine whether or not either set of measures (or some unique combination of measures from either or both sets) would differentiate good from poor readers.

In order to determine the nature and degree of these relations groups of good and poor readers were established with tasks similar to those described in semantic processing research. Four item types, presented as true-false questions, were posed following the reading of a brief paragraph. The item types included: integration of information from sentences within the paragraph; recognition of the superordinate for a noun, neither of which was in the paragraph (unprimed relationship); recognition of a superordinate for a noun included in the paragraph (primed relationship); and recognition of a sentence from the paragraph (verbatim recognition).

Method

Subjects

Sixty community college students from a two-year suburban Minneapolis college served as the subjects for the study. All students enrolled in reading and study skills classes were given the *McGraw-Hill Basic Skills System Reading Test, Form B* (Raygor, 1970). A random sample of 30 students achieving a total percentile score of 73 or above (good readers) and 30 students achieving a total percentile score of 27 or below (poor readers) on two-year college norms were selected for inclusion in the study.

Materials

A measure to determine levels of information processing consisted of thirty experimenter-constructed paragraphs. They were constructed from the category-prototype lists of Rosch (1975).

For each prototype a three-sentence paragraph was constructed. One sentence presented the prototype plus a descriptor, e.g., "The suit was denim." A second sentence presented the prototype in relation to another noun, e.g., "The model wore the suit." In half of these sentences the prototype appeared as the subject; in half the prototype appeared in the predicate. A third sentence consisted of the related noun plus a descriptor for that noun, e.g., "The model was skinny." All sentences were in the active voice. The order of the sentences within each paragraph was randomized.

For each paragraph four test items were constructed. Each was a sentence presented following the paragraph and required a true or false response. The first sentence, based on the work of Bransford and Franks (1971), consisted of the subject and the related noun plus one of the two descriptors. The subject descriptor was used half the time, e.g., "The skinny model wore the suit;" the predicate descriptor the other half, e.g., "The model wore the denim suit." In the case of false items (10 of 30) the descriptor preceded the incorrect noun, e.g., "The model wore the skinny suit."

The unprimed relationship item, based on the research of Collins and Quillian (1969) and Conrad (1972) consisted of a category-prototype statement, that is, the category and prototype were not mentioned in the paragraph. In the case of a true item the prototype and category matched, e.g., "An apple is a fruit." In the case of a false item, prototype and category did not match, e.g., "An apple is a vehicle." For both true and false items the category and prototype were other than those serving as the basis for the paragraph.

A third test sentence expanded Rosch's (1975) use of the prime in that the category-prototype statement included the prototype discussed in the paragraph, e.g., "A suit is clothing." For false items (10 of 30) the category was one of the other nine indicated by Rosch, e.g., "A suit is a vegetable."

The fourth test sentence was a verbatim recognition sentence. Ten of each of the three sentence types were selected. For true items the sentence consisted of a verbatim statement from the paragraph, e.g., "The suit was denim." For false items the descriptor was changed for a noun-descriptor sentence, and for the noun-direct object sentences a descriptor was placed before the incorrect noun, e.g., "The suit was skinny."

Order of paragraph presentation and item types within paragraphs was randomly presented to paired good and poor readers.

Procedures

An individual student, seated before a rear-screen viewer was shown the short three-sentence paragraph. When the student felt he/she understood the content of the paragraph, he/she pressed a single button causing the screen to go blank. Four items, representing the four semantic verification tasks, were then presented one at a time. To each the student responded "true" or "false" and simultaneously pressed the button. Students were then asked to indicate their confidence of response on a scale of 5 (high) to 1 (low). After the four item responses another paragraph was presented until thirty paragraphs and item sets had been presented. For both paragraphs and items presentation of the slide activated a timer that was stopped by the student's press of the button. The examiner recorded the paragraph or item reading time after each slide as well as "true" or "false" response and confidence rating for the individual items.

Design

For purposes of analysis, the data were organized according to different layouts appropriate for the several analytical procedures used. Multivariate and univariate analysis of covariance and analysis of variance for main, interaction and simple effects as well as *t* tests of group means were computed. These analyses are reported in detail by Hanson (1977). For this presentation only a subset of the results are included.

The dependent variables considered in this study were number of correct responses on the semantic verification instrument, confidence for correctness of responses, and latency of correct responses covaried by paragraph reading time on the semantic verification task.

Analyses for the dependent variables included the following: ranking and *t* tests for mean correctness and mean confidence on the semantic verification tasks; analysis of variance on correctness of response for item type by true-false by ability; and for latency of response, multivariate and univariate analyses of covariance utilizing paragraph reading time on the semantic verification tasks as the covariate.

As the semantic verification instrument was experimenter-constructed, instrument characteristics were determined, including an application of the Kuder-Richardson$_{20}$ for each of the four 30-point scales representing an item type as well as for the total 120 item test. The analysis indicated that the range of item difficulty was from .857 to .987; the reliability of subtest scores ranged from .655 to .429 and the overall text exhibited a reliability coefficient of .813.

Since the students were able to control viewing time for both the paragraph and the four item types that followed, and since items for the semantic verification task were drawn from high "goodness of example" category-prototype relationships identified by Rosch (1975) and the sentence and paragraph structure highly restricted, it was expected that verification of the truth or falsity of an item would not be difficult. In order to determine if this were the case, two analyses of the data were initiated: the first consisted of calculating mean number correct, standard deviations, ranks, and *t* tests between ability groups for mean number correct for the eight treatment conditions; the second analysis consisted of computing percentage of correct answers to true and false item types for each treatment condition. This was done as there was a possibility of 20 correct true responses but only 10 correct false responses. These converted scores made possible calculation of analysis of variance to determine if interaction existed among item type, true-false conditions, and reading ability.

Results

The first analysis, comparison of means, standard deviations, rankings, and *t* values for good and poor readers' correct true, correct false, and total correct responses indicated that good readers correctly answered significantly more items, $t(58) = 3.94$, $p < .001$, despite all subjects' ability to control both the paragraph and item viewing time. Total means were 113.93 and 108.47 for the two groups. Furthermore, examination of raw scores revealed that good readers had a higher mean number of correct responses under all eight treatment conditions.

For true items good readers had significantly more correct items for three item types: integration, $t(58) = 2.89$, $p < .005$, primed category-prototype, $t(58) = 3.02$, $p < .004$, and verbatim recognition, $t(58) = 3.68$, $p = .001$, items that required reading, or were at least influenced by the reading situation. The means for the three item types were 26.77, 24.67; 29.47, 28.53; and 27.93, 25.80, respectively.

Only one item type, verbatim recognition, favored good readers, $t(58) = 2.45$, $p < .017$, when the dependent measure was correct identification of false items. The means were 8.57 and 7.57. The lack of discrimination from the false items may have been a result of the nonsensical nature of some of them; that is, some of the items were so absurd that most subjects could reject them on the basis of logic rather than the passage in which they were embedded.

Ranking of true, false, and item type totals revealed that both good and poor readers portrayed identical rankings within item types. Correct items were rank-ordered for both groups from most correct to least correct as follows: unprimed category-prototype relationships, prime category-prototype relationships, verbatim recognition, and integration.

To consider how item responses varied in terms of the ability groups, true as compared to false items, and item types, and to determine if interaction between these factors existed, analysis of variance was carried out on percentage conversions of the raw scores. One between factor, ability, and two within factors, true versus false, and the four item types were considered.

Results indicated that while all three main effects achieved significance at the .01 level, there were also two interactions that were notable: ability by item type, $F(3,147) = 3.89$, $p < .01$, and true-false by item type, $F(3,174) = 27.92$, $p < .01$. The first interaction demonstrated that while reading-required item types were more difficult than non-reading-required item types for both ability groups, they were a great deal more difficult for poor readers. However, the second interaction between true-false conditions and item type complicated the interpretation of the main effects previously indicated in that most of the variation in these two factors stemmed from two extremely depressed cell means: false verbatim recognition and false integration items.

The percentage of errors for totaled item types was 5.05% for good readers and 9.16% for poor readers. Some subjects made no errors on certain item types resulting in insufficient data on which to make judgements regarding incorrect responses. The attention to correctness through trials, and request for confidence rating of responses seemed to have effectively encouraged subjects to concentrate on correctness. For these reasons analysis of confidence ratings and latencies focused on correct items only.

Despite the ability of subjects to control the length of time they had to view the paragraphs and the items, differences in response correctness were noted on particular items. In order to determine if these differences might be a result of overreaction or a thoughtless response, subjects were asked to indicate how confident they were of their correctness.

For both good and poor readers there was a high degree of confidence for their correct responses, all confidence means being at 4.61 or above. Ranking the mean true and mean false confidence indicated that for both good and poor readers the order of highest to lowest confidence was identical to that for greatest to least in correctness of response.

To determine differences in processing, mean latencies for the eight treatment conditions for each ability group were calculated and, as would be expected, good readers responded more rapidly in completing the reading of the paragraphs and in correctly verifying items under the eight treatment conditions.

To attenuate reading speed as a confounding variable in the consideration of the data, subjects' response times to the items were covaried by paragraph reading time and the resulting times considered by multivariate analysis of covariance which simultaneously compared all four response times of good and poor readers to determine if differences did exist between the two groups for the adjusted latency measure.

The adjusted means resulting from the multivariate analysis of covariance, reflect latencies that significantly favor good readers for both true, $F(4.54) = 4.10$, $p < .006$, and false, $F(4.54) = 4.80$, $p < .002$, items. Not only were good readers faster than poor readers in their overall processing speed, but they also appeared to be faster in the completion of semantic verification tasks.

A second univariate analysis of covariance of latencies compared good and poor readers on each of the eight treatment conditions (four item types, true-false). Good readers responded more rapidly to semantic verification tasks in seven of the eight treatment combinations of levels; the one exception was false verbatim recogniton.

Conclusions

As the attempt of this study was to examine relationships between two fields of study whose common bases had not yet been established, limitations must be recognized in interpreting the results. Certain tentative conclusions, however, are warranted.

Good adult readers do exceed poor adult readers in their ability to respond correctly to semantic processing tasks that require comprehension of written discourse. It would appear that good readers readily integrate new learnings acquired through reading into their conceptual framework while poor readers deal with the reading task as a singular presentation of information unrelated to what they already know.

In terms of the item types themselves, true items favored good readers for three of the four tasks, the exception being that of unprimed category-prototype relationships. The differences that were found were not a result of confusion or lack of attention to the task. This was established by the few differences in confidence for correct responses by item type.

Analyses of the latency scores for correct responses utilizing paragraph reading time as a covariate indicated that good readers responded to both true and false items more rapidly than poor readers. Additional analyses, again using paragraph reading time as a covariate, demonstrated that good readers responded more rapidly in terms of semantic verification to seven of the eight item type true-false tasks.

These findings demonstrated that there were statistically significant differences by good readers on a majority of the semantic verification tasks. It would appear that the developmental differences in semantic processing favoring good readers, established by several writers, have been substantiated with adult subjects by the results of this study.

References

Battig, W. F. & Montague, W. E. Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology*, 1969, *80*, 1-46.

Bransford, J. D., Barclay, J. R. & Franks, J. J. Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology*, 1972, *3*, 193-209.

Bransford, J. C. & Franks, J. J. The abstraction of linguistic ideas. *Cognitive Psychology*, 1971, *2*, 331-350.

Collins, A. M. & Quillian, M. R. Does category size affect categorization time? *Journal of Verbal Learning and Verbal Behavior*, 1970, *9*, 432-438.

Collins, A. M. & Quillian, M. R. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1969, *8*, 240-247.

Conrad, C. Cognitive economy in semantic memory. *Journal of Experimental Psychology*, 1972, *92*, 149-154.

Fillmore, C. J. The case for case. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory*. New York: Holt, Rinehart and Winston, 1968.

Freedman J. L. & Loftus, E. F. Retrieval of words from long-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1971, *10*, 107-115.

Firdja, N. H. Simulation of human long-term memory. *Psychological Bulletin*, 1972, *69*, 1-31.

Grober, E. H. & Loftus, E. F. Semantic memory: Searching for attributes vs searching for names. *Memory and Cognition*, 1974, *2*, 413-416.

Hanson, K. M. *The Relationships Between Reading Ability and Semantic Verification Tasks.* Unpublished doctoral dissertation, University of Minnesota, 1977.

Loftus, E. F. Category dominance, instance dominance, and categorization time. *Journal of Experimental Psychology*, 1973, *97*, 70-74.

Loftus, E. F. & Grober, E. H. Retrieval from semantic memory by young children. *Developmental Psychology*, 1973, *8*, 310.

Meyer, D. On the representation and retrieval of stored semantic information. *Cognitive Psychoogy*, 1970, *1*, 242-300.

Minsky, M. (Ed.). *Semantic information processing.* Cambridge, Mass.: M.I.T. Press, 1968.

Quillan, M. R. Semantic memory. In M. Minsky (Ed.), *Semantic information processing.* Cambridge, Mass.: M.I.T. Press, 1968.

Raygor, A. L. *McGraw-Hill Basic Skills System Reading Test, Form B.* Monterey, Calif.: CBT/McGraw-Hill, 1970.

Rosch, E. Cognitive representations of semantic categories. *Journal of Experimental Psychology*, 1975, *104*, 192-233.

Rosch, E. H. On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language*. New York: Academic Press, 1973.

Rumelhart, D. E., Lindsy, P. H. & Norman, D. A. A process for long-term memory. In E. Tulving and W. Donaldson (Eds.), *Organization of memory*. New York: Academic Press, 1972.

Schaeffer, B., Lewis, J. A. & Van DeCar, A. The growth of children's semantic memory: Semantic elements. *Journal of Experimental Child Psychology*, 1971, *11*, 296-309.

MICHAEL L. KAMIL
RAYMOND H. HANSON
Purdue University

Perceptual versus semantic information processing in semantic category decisions

Rosch (1975) has developed a paradigm for examining the use of advance information in making semantic category decisions. Two features of this paradigm are worth discussing in detail. First, Rosch has developed a set of norms for "level-of-goodness" for the semantic categories she uses. That is, each category is represented by exemplars which have been rated on the basis of how well they fit the categories. For example, in the category 'FRUIT', 'apple' is a high level-of-goodness exemplar while 'kumquat' is a low level-of-goodness exemplar. The level-of-goodness norms seem to be stable and reliable as indicated both by Rosch's work and that of Childrey and Kamil (1977).

Second, Rosch uses a decision task which incorporates advanced information, known as "priming". The decision task is to make "same category" or "different category" judgments for pairs of exemplars from the categories. On half of the trials, advanced information is provided in the form of the name of a category relevant to the word pair. On the remainder of the trials, no priming information is provided, in which case the test pair is preceded by the word 'BLANK' instead of a category name. Three types of word pairs are included, physically identical pairs, categorically identical pairs, and different category pairs. Physically identical trials contain the same word, repeated, for example, 'Apple ** Apple'. Categoricaly identical pairs have two different words taken from the same category, for example, 'Apple ** Orange'. Different pairs have two different words taken from different categories, at the same approximate level-of-goodness, for example, 'Car ** Apple'. Subjects respond "*same*" to both physically identical and categorically identical pairs by pressing one key and "*different*" to the different category pairs by pressing another key. The dependent measure was reaction time (RT) to press the correct key.

Rosch found that level-of-goodness was related to latency of decision for categorically identical pairs with high level-of-goodness producing faster RTs. Priming also produced faster RTs for categorically identical pairs. However, when the pairs were physically identical, priming interacted with level-of-goodness. These results have been replicated and extended by Hanson & Kamil (1977).

The explanation Rosch applied to these results was that subjects generated some "aspect of meaning" on primed trials. A semantic aspect of meaning uniformly facilitated processing of primed categorically identical pairs as a function of level of goodness. For physically identical pairs, the interaction occurs because the perceptual aspect of meaning is more like good exemplars (facilitating decisions) than poor exemplars (inhibiting decisions).

Hanson and Kamil (1977) attempted to replicate the Rosch study. They found differences in processing of physically identical pairs, dependent on the modality in which advanced information was presented. In the original Rosch study, primes were presented auditorily. When primes were presented visually in the Hanson and Kamil study, the interaction for physically identical (PI) pairs was not obtained. These results indicate that the facilitation and inhibition of different level-of-goodness pairs involve the use of visual centers. The PI results can be understood as a visual biasing in favor of a few, highly probably visual stimuli and against others.

This visual expectancy is disrupted by the use of visual advanced information in a way similar to modality specific interference (Brooks, 1967, 1968). Processing the prime word visually prevents the use of those visual centers for simultaneously preparing for a few, likely stimuli. Hence visual priming results in no interaction for physically identical pairs. In either modality, the priming effect is unaltered for semantic (categorically identical) decisions. This indicates that these semantic effects are "automatic". In contrast, the perceptual effects observed for physicaly identical pairs are easily disrupted. This suggests that the PI effects result from perceptual strategies which subjects adopt to facilitate processing (cf. Hanson, Kamil & Snyder, 1978).

Method

Sixteen junior high school students participated in the study. Eight were selected from remedial reading classes and eight were selected from non-remedial English classes. All students were paid for participation in the enrollment. Sixteen college students also participated. They received course credit.

The independent variables in the normal Rosch paradigm are: 1) Type of trial (physically identical, categorically identical, or different); 2) Level-of-goodness (high, medium or low); and 3) Advanced information (primed or unprimed). In this study, modality of advanced information (visual or auditory) was also manipulated. Finally, reading ability (college, good or poor) was another variable. In all cases, the dependent variable is latency of decision.

Experimental sessions lasted about one hour. Each participant received 196 trials, half primed and half "blank", equally divided between "same" and "different". In the present study, stimuli were presented on a cathode ray tube display controlled by an ADS 1800E computer. Reaction times to make decisions were measured and recorded automatically.

Results

All data were subjected to repeated measures analyses of variance. Where appropriate, quasi *F*-ratios were constructed. In the interest of readability, the precise *F* values have been omitted. In all cases, α was 0.05 or smaller.

Figure 1 shows the results for college students, good junior high school readers, and poor junior high school readers. College students show a level-of-goodness effect in all conditions. Good readers show a level-of-goodness effect in all conditions except in the physically identical trials with visual primes. Poor readers show no level-of-goodness effects in any conditions.

For the categorically identical trials, college students show priming advantages. Good readers show priming advantage *only* when the priming information is auditory. Poor readers do not show any advantage of priming information.

For the physically identical trials, college students show an interaction between priming and level-of-goodness for auditory information. Neither good nor poor readers show the interaction effect.

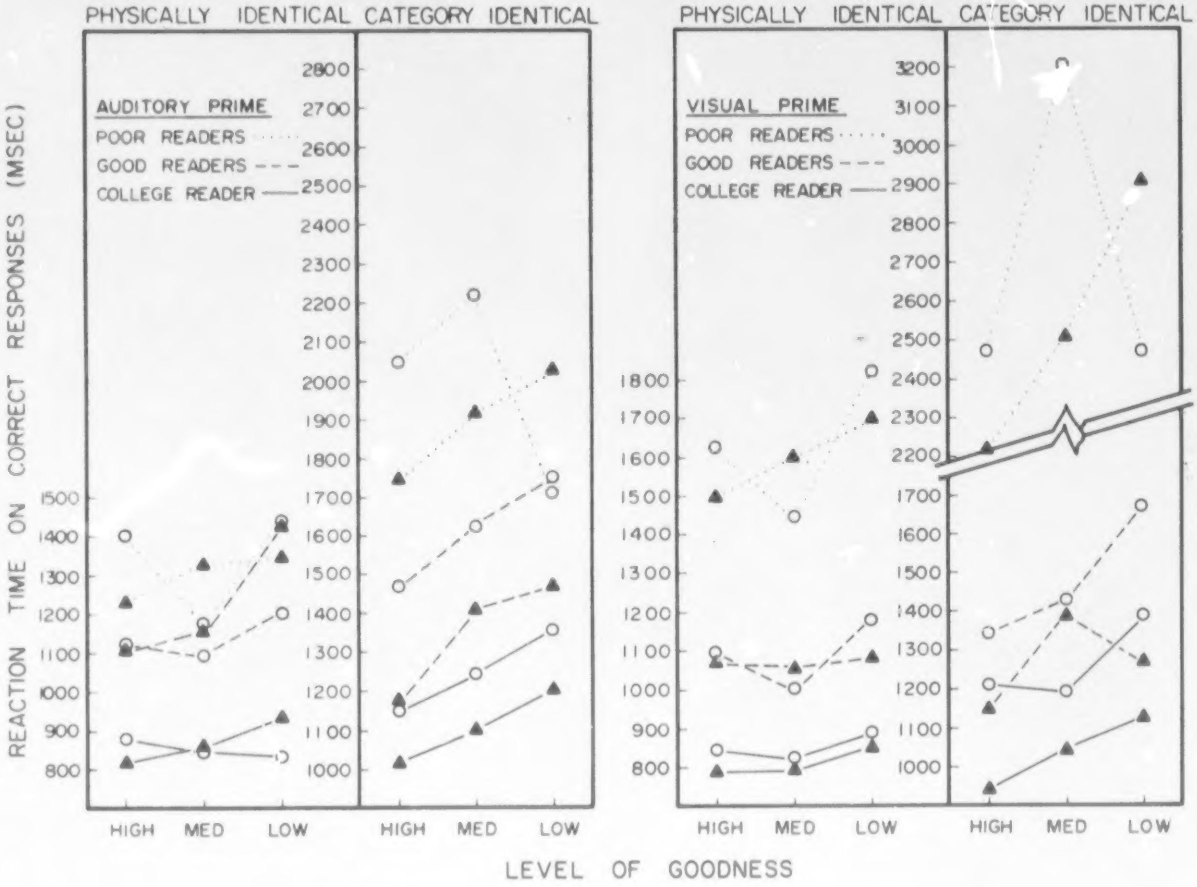
Finally, college students are faster than junior high school students, and good readers are faster than poor readers.

Discussion

These results reveal some interesting "developmental" trends over the range of students tested. Examination of Figure 1 makes these trends most apparent.

95

Figure 1. Reaction times as a function of type of trial, modality of prime, level-of-goodness of exemplar and reading level.



Within the interpretive framework of semantic (automatic) and perceptual (strategy) effects, college students and good readers show similarities for semantic effects in both auditory and visual conditions. This suggests that both of these groups have developed a well-organized semantic memory network which is reflected by level-of-goodness norms and decision latencies. Of the two modality conditions, however, the auditory condition is easier for good readers and the results for this group match college students more closely. On physically identical trials, the auditory condition shows good readers developing the necessary level-of-goodness effect. However, they do *not* seem to differentiate their perceptual strategy for primed and unprimed trials to yield the physically identical interaction.

Poor readers show neither the perceptual nor the semantic effects. Thus, the major implication of the present data is for a general trend in reading development. First a stable semantic category structure is established. Then, by college level, perceptual strategies based on that structure are developed. Junior high school students are still developing the category structure (in terms of level-of-goodness) found in college students. The poor readers clearly showed reaction times which were very slow compared to good readers. They were slower still for visual primes than for auditory primes. Good readers are further along in reading develoment than are poor readers. Specifically, good readers seem to be better than poor readers at using advanced information for facilitating semantic decisions and adopting perceptual strategies.

While the Rosch task used in this study does not involve the entire spectrum of reading skills, it does isolate and examine two skills important in reading: semantic and perceptual processing. Since poor readers have difficulty with this relatively simple semantic task, it would be expected that they would have even more difficulty with the more complex semantic demands of a reading task. Similarly poor readers have more difficulty than good readers in the perceptual aspects of this task though both groups are poorer than college students. Because perceptual strategies are based on a stable semantic structure, it may be that junior high school students need development of their semantic structures and their consequent perceptual strategies. Additional work will be needed to determine 1) whether these findings are true when larger units (than words) are used and 2) whether these differences found between good and poor readers can be remediated by conventional methods.

References

Brooks, L. R. The suppression of visualization by reading. *Quarterly Journal of Experimental Psychology*, 1967, *19*, 289-299.

Brooks, L. R. Spatial and verbal components of the act of recall. *Canadian Journal of Psychology*, 1968, *22*, 349-368.

Childrey, J. A. & Kamil, M. L. *A preliminary examination of the relationship between category structure and reading.* Paper presented at the annual meeting of the National Reading Conference, New Orleans, December 1977.

Hanson, R. H. & Kamil, M. L. *Modality specific effects in semantic category matching.* Paper presented to the Midwest Psychological Association, Chicago, May 1977.

Hanson, R. H., Kamil, M. L. & Snyder, C. R. R. *Attentional versus automatic processing in semantic category decisions.* Presented to the Midwestern Psychological Association, Chicago, May 1978.

Rosch, E. Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 1975, *104*, 192-233.

PRISCILLA A. DRUM
University of California, Santa Barbara

Prose recall responses and categories for scoring

The recall of what has been read contains responses that closely approximate the original and responses that vary on some undefined continuum of difference from the original. Much of the experimental research on reading comprehension indicates that competent readers recall the substance of the passage, e.g., the text base (Kintsch, 1975, 1976), the redundant information (Drum, 1974, 1976), the hierarchical structure (Meyer & McConkie, 1973; Filby, 1975), and the story structure (Thorndyke, 1977). These studies used various analysis systems for determining the structure of the passages; the systems then become templates for scoring the recall responses since the major interest or outcome was to determine the correspondence between the passages and the recall.

However, the constructive view of memory (Bransford & Franks, 1973) postulates an integration of fragments within a text with a plan for what must have occurred based on generalized knowledge of the word. Interference can come from inadequate or inappropriately used knowledge or from selecting nonessential text fragments. Deviations from the text do occur even with college undergraduates, the typical subject for these studies. Recall for "The War of the Ghosts" (Bartlett, 1932) and for the biblical story "Joseph and His Brothers" (Kintsch, 1976) include content additions, which were explained as prefamiliarity with the "Joseph" story or as fitting the "Ghost" information into expectations by the reader for a cultural norm. These text deviations and some of the generated responses in the Cofer (1973) study seemed reminiscent of the classroom recall of young children after reading a story. Text recall is not just the reader's creation, but an adaptation of the information in the text to his own prior memorial structure for similar information. Responses are interactions between what was known before reading and what the text said.

A purpose for this study is to describe the difference response. The first step is to analyze text from different responses. To do so, the original passages are analyzed into a hierarchy of propositions using procedures similar to Kintsch's (1975) for deriving the text base. This consists of "a structured list of *propositions*. A proposition consists of a predicate with one or more *arguments*; arguments are concepts or propositions themselves" (Kintsch, 1977). The attempt is to delineate the meaning relationships of the text which can be realized with different function words and different orderings — i.e., active to passive alterations, or progressive to active changes. The linear ordering of propositions in a hierarchy established by argument repetition describes *a* structure of the passage, which is used as a template in scoring textual responses that adhere to the meaning but deviate somewhat in syntactic and function word expression.

The protocols are also analyzed into propositions where possible. There are often parenthetical remarks and incipient phrases which cannot constitute a proposition and, therefore, a word count rather than a proposition count is used. Also, since the subjects do not have time for editing and revision, their hierarchical structures are not described.

Four ordinal categories have been developed to encode responses on a continuum from text specific to text external. The categories are global in that each contains subcategories with obvious explanatory power which are summed together. The categories are: 1) *Text specific responses* — The units scored as text specific must correspond with the text propositions unambiguously. They are restatements of text propositions, including acceptable syntactic paraphrase, synonymy of elements with the same conceptual referent, and substitutions of pronouns appropriately referenced elsewhere. 2) *Text entailed responses* — The units scored as text entailed include text specific elements put together in new ways and additions of information that are semantically entailed by the text and are correct according to a content expert. 3) *Text evoked responses* — The units scored as text evoked also include elements of the text, but they are either inappropriate recombinations, additions of information external to text elements, or general statements that do not convey any specific information. 4) *Text external responses* — The units scored as text external either have no relationship to the text, such as storytelling conventions, self-report comments, or repetitions of previously recalled statements.

The four categories and the procedures for scoring are described in more detail in Drum and Lantaff (1977).

Method

Sixteen eighth grade pupils, half of them able readers and half below average in reading ability, read four passages. Ability was determined by scores at the 90th percentile or above or below the 50th percentile on the SRA, Form E, reading comprehension section. The subjects also represented equal numbers of boys and girls. After reading a passage, each student was asked to orally recall the information. A day later, delayed recall for two passages was taken; a week later, the remaining two. The cue for delayed recall was the one-word title provided with each passage. All oral responses were taped. Each subject knew in advance that he would be asked for recall of the information by title both immediately after reading and either a day or week later. They were also told to read as if they were reading a school assignment. Nobody was apparently attempting to rote memorize the material. The recorded reading times averaged 86 seconds, too brief for attempted memorization of the passages.

The four passages (252-275 words) were selected from science and social studies texts. The Fry readability levels for the passages were from 7.4 to 8.0. The passages were analyzed into propositional hierarchies — four to seven levels in the hierarchies and 123 to 147 propositions.

The design was a repeated measures, mixed factorial with ability and sex between subject variables and type of passage, passage and amount of delay within subject treatment factors.

The immediate and delayed recall responses were analyzed into propositions, with function words counted as they appeared in the propositions. Every propositional unit was then given one of the four category designations. Each recall response was scored by three people. There was substantial agreement among scorers, approximately .80; above .95 for text and external and .60 to .70 for entailed and evoked. Any disagreement about a unit was reanalyzed until an agreement could be reached. Then the words per category per passage became the scores for further analysis.

The significance of the category distinctions by ability, by sex, by type passage, by particular passage within type, and by amount of delay were tested using analysis of variance procedures.

Table 1

Means and Standard Deviations by Ability for Protocol Categories

	Total		Text		Entailed		Evoked		External		
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	(N)
At Immediate Recall											
Ability:											
High	141	50	50	26	49	20	18	15	23	14	(32)
Low	87	37	17	15	26	11	26	17	18	11	(32)
At Delayed Recall											
High	115	43	34	19	45	21	17	10	19	12	(32)
Low	69	33	8	6	26	18	22	13	13	8	(32)

Results & Discussion

The major category means indicate that able readers remember more text responses at both immediate and delayed recall, $F(1, 12) = 9.8, 12.2, p < .01$, and more entailed responses at immediate recall, $F(1, 12) = 7.5, p < .05$. Less able readers have more evoked responses, particularly for science passages, at immediate recall test point. Delayed recall has on the average fewer propositions stated, but what is retained is generally similar to immediate recall. The proportionate stability of entailed and evoked responses from immediate to delayed recall as compared with the decrease in textual recall indicates that reconstructed responses are more likely to be remembered over time.

The ability by type of passage interaction for evoked responses, $F(1, 12) = 4.9, p < .05$, for immediate recall is due to an increase in general responses by less able readers for the science passages, science mean of 28 versus social studies mean of 23, while able readers have fewer evoked responses for science (mean of 16) than for social studies (mean of 20). Less able readers are less likely to make delayed entailed responses for science passages (a mean of 13) than for social studies (a mean of 38), $F(1, 12) = 17.6, p < .01$. Able readers do not differ. The only other interaction for ability is less able readers make fewer external comments for the two passages "Engine" and "Women;" they also have fewer text and entailed remarks and more evoked for these passages. However, able readers also make fewer text and more evoked for the same two passages, but they also increase their external comments, $F\ (1, 12) = 11.9, p < .01$. Able readers may be simply inclined to make more statements. If they do not remember much of the passage, they are likely to make more comments about their poor performance.

Science passages were more difficult for these children at both test points, less text responses for all readers, $F\ (1, 12) = 27.0, 17.3, p < .01$. Though readability was approximately the same for all four passages, reading and life experience with methods for conserving heat in buildings and for the way gears work in operating machinery is likely to be less for eighth graders than descriptions of the household work of women in the frontier or the problems of an explorer in the new world. The science passages are technical, providing information about inanimate things. The social studies passages are about people. Human actions and reactions are familiar events, and readers can use their prior knowledge about people to structure the information and thereby remember it.

The results for "Engine" exemplify the possible effects of experience on learning from reading. More boys and girls may have tinkered with gears, not an innate predisposition but a cultural phenomenon. Though not significant, girls of both ability groups generally have more text specific and less evoked responses. For "Engine" this pattern is reversed. Text entailed responses are identical for able readers generally, while low girls have more of such responses than low boys. For "Engine" again the pattern is somewhat reversed; boys of both ability groups generally out-perform the girls in the respective groups. In delayed recall the major change is more entailed responses by able boys versus able girls, a switch from their general performances.

"Engine" is the most technical article. It requires more prior experience to comprehend. There are phrases with two underlying meanings: for example, the *crankshaft gear* can mean *it is a gear* or *it has a gear*. General reading ability is the most potent factor in learning from text, but experience also makes a contribution; at least, experience is inferred to be the source for the differences in girls' high performance generally and relatively low performance on "Engine."

Explorations into the interactions between experience and the content of material are indicated. Blocking by prior knowledge of the content as well as by reading ability and examining both the text and the difference responses may aid in determining the components of learning from text and the processes involved in reading comprehension.

The categorical analysis applied to subject recall responses does indicate differences between good and poor comprehenders. Though each student's responses varied as a result of the different passages, the categorical patterns of more text and entailed for an able reader remained constant over passages. Categorical difference scores for each subject for each passage were also calculated. Able readers were more likely to bring together information from all parts of the passage, to make text derived inferences, and to add case-linked arguments. Less able readers seemed to divide into two groups: those that stated a text proposition and then repeated it, or those who picked up one idea and then elaborated on it though there was no apparent relationship to the text in the elaboration. Since all of the passages had a readability close to grade level, one would have to see if the recall patterns held for simpler materials for the less able and harder materials for the good readers. If replicable and replicated, categorical analysis could help define what is meant by reading comprehension.

References

Barlett, F. C. *Remembering*. Cambridge: Cambridge University Press, 1932.

Bransford, J. D. & Franks, J. J. The abstraction of linguistic ideas. *Cognitive Psychology*, 1971, *2*, 231-350.

Cofer, C. N. Constructive processes in memory. *American Scientist*, 1973, *61*, 537-543.

Drum, P. A. *Between sentence factors and the effects on reading rate and recall*. Unpublished dissertation, Stanford University, 1974.

Drum, P. A. & DeFratis, E. *Retention of information presented within school texts*. Paper presented at the International Reading Association National Conference, Anaheim, May 1976.

Drum, P. A. & Lantaff, R. E. Scoring categories for protocols. Paper presented at the Second Annual Language Conference, Boston University, October 1977. (In press, *Conference Proceedings*.)

Kintsch, W. *The representation of meaning in memory*. Hillsdale, New Jersey: L. E. Erlbaum Associates, 1975.

Kintsch, W. Memory for prose. In C. N. Cofer (Ed.), *The structure of human memory*. San Francisco: W. H. Freeman & Co., 1976.

Kintsch, W. & Vipond, D. Reading comprehension and readability in educational practice and psychological theory. Paper presented at the conference on memory, University of Uppsala, June 1977. (In press, *Conference Proceedings*).

Meyer, B. J. F. & McConkie, G. W. What is recalled after hearing a passage? *Journal of Educational Psychology*, 1973, *65*, 109-117.

Thorndyke, P. W. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 1977, *9*, 77-110.

ROGER E. LANTAFF
University of California, Santa Barbara

Two factors affecting text recall

A useful approach in discussing reading comprehension is to view it as a process of acquiring information. During this process, the reader utilizes surface and base characteristics of the text as well as his internalized psychological processes and existent knowledge in deriving a meaning from this written text (Kintsch & Vipond, 1977). Mature readers reconstruct their version of the author's intended message in part by utilizing the text as a kind of processing plan (Freedle & Carroll, 1972) and then integrating or assimilating this acquired information to existing cognitive schemata.

According to the "levels of processing" framework, proposed by Craik and Lockhart (1972), the nature and duration of a memory trace is determined by the kind of processing that is performed on input; an analysis of the recall that readers produce for texts they have read does in many ways seem to reflect the processing that has occurred. The versions readers reconstruct usually include information that is specific to the text and inferences that were implied but not explicitly stated in the text, as well as information conceptually related to the text that seems to intrude into the subject's recall from existing knowledge. Consequently, it is suggested that a close look at the recall characteristics for differing passages may provide clues which enable us to make inferences about the processes of comprehension as well as the text variables which influence this processing.

This study examines the nature of text recall obtained at different points in time for passages varied by subject and reading difficulty. In additon, a nested condition investigates the effect of delaying initial recall on subsequent recall.

Method

Subjects

The subjects for this study were 16 graduate students from the University of California, Santa Barbara Graduate School of Education.

Materials

Five passages were selected from science or social studies texts, designed for fifth grade students and college undergraduates. These were texts currently available for use in the public schools. Each selection was approximately 250 words long with a readability between 5.0 and 5.8 for the easy, and 12.8 and college for the difficult passage conditions: a hard science, a hard social studies, an easy science, and an easy social studies. The easy social studies condition contained two passages because of the nested condition. The order of the four conditions was systematically varied for presentation with the position of the two easy social studies passages counterbalanced within the design so that the passage which was *not* to be recalled at the immediate recall session always occurred in the fifth position.

Text Analysis

Stimulus passages were selected and then decomposed sentence by sentence into the clausal units contained within. These clausal units or "clausal propositions" were comprised of a predicate and n-tuples of arguments. These functioned as constituents of a subject or predicate term. A distinguishing characteristic of these "clausal propositions" was that their predicate term was marked by a grammatical verb in the text's surface structure.

Once the sentence was parsed into clauses, these were arranged hierarchically according to a specified set of rules for assessing clause relations, based primarily upon the sentence-internal coordination, subordination or complement relationships and levels of argument generality. After the first sentence was so arranged, the next sentence was analyzed in a similar manner. These clusters of propositions were then connected on the basis of the relationship between the head or nucleus (Longacre, 1977) proposition for each cluster. If the head of the second cluster acted to further describe, specify or elaborate on the information of the first, it and all its dependent or marginal propositions assumed a lower position in the hierarchy. If, on the other hand, this second proposition introduced essentially new information that functioned as topical or thematic development, as opposed to propositional elaboration, then the head propositions assumed positions on the same hierarchical level. Each successive sentence was then analyzed and arranged accordingly.

Procedure

Each of the passages selected was typed and randomly assigned to the students at the initial reading session. They were instructed to read each passage twice, first skimming it to get the general idea and then going back to read the passage more carefully. They were asked to try to read these passages in the same way they would for a class assignment. In addition, they were asked to record the times they began and finished reading.

When the subjects finished reading a story, they turned the passage over and, on a separate sheet which contained only the title of the preceding story, wrote down everything that they could remember. The subjects had unlimited time for recall. Following the reading of the fifth passage, the subjects were informed that no recall on this passage was required.

Forty-eight hours later, the group was again asked to recall the passages. For this session, as well as the third, the subjects were asked to recall all five of the passages. The third and final recall session took place one week after the initial reading session. Unlike the previous sessions in which subjects were furnished only with the passage title as a prompt, the "one week" recall sheets each contained four noun "cues" selected from the appropriate stimulus passage. Subjects were requested to recall all that they could about the cues and then to add whatever else they could remember about the story. This produced three recall conditions: an immediate free, 48-hour free, and a one-week cued recall condition.

Following these, the protocols were analyzed in a manner similar to the texts and partitioned into lists of clausal propositions. These were not arranged hierarchically, however. Each of these recall propositions was then compared to the corresponding text list and classified as belonging to one of three response categories. (There are actually five categories in our analysis [see Drum & Lantaff, 1977], but two did not prove appropriate for college students with written protocols.) The recall was scored as either *text specific* (verbatim propositions, allowing for synonym substitution); *text entailed* (propositions which are bounded by the text, containing specific arguments and predicates arranged in new ways); or *text evoked* (instrusions and generalizations).

Results and Discussion

A summary of recall category means is presented in Table 1.

Table 1

Means and Standard Deviations for Response Categories by Session

Response Categories:	Total	Text-specific	Entailed	Evoked
Session 1	16.58(6.49)	9.66(6.05)	6.38(2.60)	.54(.97)
Session 2	11.92(5.77)	5.22(4.01)	5.81(3.08)	.91(1.18)
Session 3	11.33(5.49)	4.20(3.48)	6.09(2.84)	1.03(1.26)

As might be expected, there is a distinct drop in the total number of clausal propositions remembered between the immediate and 48-hour conditions, with the cues (four nouns) acting to stabilize recall between 48-hours and a week. There are also qualitative changes that occur in the types of information recalled. There is essentially a trade off that occurs between the text specific and text entailed categories; whereas the amount of specific recall diminishes, the amount of entailed recall shows proportionate gain. It seems that while surface level pragmatic relations fade from memory fairly quickly, subjects are still able to reproduce text-relevant information. This finding would be in line with a constructivist interpretation of memory.

The low amount of evoked recall is also of interest. In previous studies, researchers have found that information seemingly related to a subject's previous experience frequently intrudes into recall protocols when the passage content is highly familiar to the subjects and/or the time increases between reading and recall (Kintsch, 1974). This was not the case for these students.

The topics of these passages were such that normal readers would be expected to have well developed conceptual bases due to prior experience. Students should, it would seem, "ad lib" quite a bit as memory for these texts fade; however, the responses were quite conservative. It seems that when the memory for text specific relations gets weak, mature readers still maintain a "sense" of the text, what Kintsch has called a "macrostrsucture". These students were reasonably sure of what they had not read and edited or constrained their recall accordingly.

Passage Differences

Analysis of variance for the total recall of the passages yielded a significant difference for reading difficulty of texts, $F(1,15) = 4.91$, $p < .05$, but not for the texts of different types (science vs. social studies). There was however a significant interaction between text type and text difficulty, $F(1,15) = 4.57$, $p < .05$. The primary factor contributing to this interaction is apparently the large recall difference between the science passages which does not occur between the social studies texts. The two social studies texts are quite distinct in terms of their measured reading difficulty and yet are remembered with close to equal effectiveness. On the other hand, the two hard texts are similar in readability scores, but are remembered unequally. This raises some question about formula-measured readability and its relationship to memorability and comprehensibility. It seems that if these formulas are measuring lexical and grammatical complexity, the frequent practice of implying that complexity equals difficulty, without considering other reader and text variables, may prove to be a precarious one.

There is a relationship between the number of propositions in a passage and the amount of recall, in which the presence of more propositions in the text base correlates with higher recall. It is difficult to believe that fewer propositions simply mean that fewer are available for recall when the totals far exceed memory limits. Rather, it seems that since the number of words in a passage is approximately the same and yet the number of propositions differs (51 for the easy science, 48 for the hard social studies, 46 for the easy social studies and 41 for the hard science passages), there is what might be called a proposition complexity factor. Propositions are composed of arguments and predicates, with the arguments serving a function in which they specify text referenced concepts. In the fifth grade passages, the arguments (noun phrases) are quite basic, usually two or three words at the most, while arguments in the college passages tend to be much more complex. This complexity or density of a proposition's constituents may influence a proposition's memorability. More likely, however, this density is an indicator of thematic specificity and the more specific the topic, the less likely the reader is to have an adequate conceptual base into which to integrate information learned from the text. This brings us to what seems a key factor in the differences in recall: prior knowledge. The hard science passage on memory was unmistakenly the most esoteric of the stimulus texts and, it is presumed, the least familiar to the subjects. The mean difference between this passage and the others, which were more general in focus and thematically more likely to be part of one's knowledge about the world, suggests a powerful influence of the reader's prior knowledge on the intergration and rearrangement of the selected text information.

Immediate vs. Delayed Recall

As mentioned, the students were presented with two easy social studies passages, counterbalanced and arranged so that one passage occured in rotation with the other three pasage types and the second always in the fifth and final reading position. As with the other three passage types, students were requested to recall the first easy social studies text immediately after reading it, whereas there was no recall required for the second passage until the 48-hour delayed recall session. This condition was designed to ascertain the effects that delaying the students' initial written recall of a passage would have on subsequent memory for that material.

While there is no statistical difference between the stories, there was a significant difference between the immediate and delayed conditions for total recall, $F(1,15) = 4.69$, $p < .05$, with total and text specific recall much higher

---

[1] Statistics reported have been corrected for possible bias due to repeated measures through use of the Geisser-Greenhouse Conservative F test.

when students recalled the text immediately than when they delayed their recall 48 hours. This was the case despite the fact that students spent a substantially longer amount of time reading the second, unrecalled passage.

Table 2

Means and Standard Deviations for the Immediate vs. Delayed* Initial Recall

Immediate

Passage 1			*Passage 2*			*Overall*		
Total	Text	Reading Time	Total	Text	Reading Time	Total	Text	Reading Time
11.88	7.00	119.38 sec.	13.38	5.50	115.00 sec.	12.63	6.25	117.19 sec.
(6.33)	(3.85)	(57.91)	(7.11)	(3.85)	(24.05)	(6.55)	(3.80)	(42.89)

Delayed

Passage 1			*Passage 2*			*Overall*		
Total	Text	Reading Time	Total	Text	Reading Time	Total	Text	Reading Time
5.25	2.88	166.25 sec.	8.00	1.50	163.13 sec.	6.63	2.19	165.00 sec.
(7.[illegible]9)	(4.55)	(120.85)	(5.95)	1.77	(84.00)	(6.48)	(3.41)	(104.25)

* for the second session

These recall differences held up for the third recall session as well. This seems to indicate that the nature of the activity which follows reading plays a substantial role in the process of assimilating to existing knowledge bases the information from reading a passage.

Conclusion

The results of this study suggest that the comprehensibility of a text, as indexed by the ability of subjects to recall its content, cannot be determined by readability characteristics alone. It seems, as Kintsch and Vipond (1977, p. 41) tell us, that "readability is not somehow an inherent property of texts, but is the result of the interaction beween a particular text (with its text characteristics) and particular readers (with their information-processing characteristics)." This study also seems to indicate the pressing need for studies which investigate how the reader's knowledge base affects the comprehension and learning processes and that a text recall paradigm is a valuable way to investigate this question.

References

Craik, F. I. M. & Lockhart, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 671-684.

Drum, P. A. & Lantaff, R. E. *Scoring categories for protocols.* Paper presented at the Boston Reading Conference, 1977 (In press, conference proceedings).

Freedle, R. O. & Carroll, J. B. Language comprehension and the acquisiton of knowledge: Reflections. In J. Carroll & R. Freedle (Eds.), *Language comprehension and the acquisition of knowledge.* Washington, D. C.: V. B. Winston, 1972.

Kintsch, W. *The representation of meaning in memory.* Hillsdale, New Jersey: L.Earlbaum Associates, 1974.

Kintsch, W. & Vipons, D. *Reading comprehension and readability in educational practice and psychological theory*. Paper presented at the Conference on Memory, University of Uppsala, 1977 (In press, conference proceedings).

Longacre, R. E. *An anatomy of speech notions*. Amsterdam: Peter de Ridder Press, 1977.

JAMES FLOOD
Boston University

A system for scoring readers' recall of propositions from texts

Text information is decoded, recoded, and encoded in the reading/recall paradigm. During recall sessions, readers generate numerous propositions that are the results of operations performed both during and after the processing of texts. Because no evidence exists at the present time to conclude that verbatim recall is equivalent with the absence of processing/comprehension during reading, it must be assumed that all recall is the function of operations performed by readers. All operations that readers perform may be referred to as inferences.

Evidence of inference that readers generate has been derived from the coding of elicited data in recall. Kintsch (1977) has devised a prototypic scoring system for coding general recall of readers. Texts and reader's recall are decomposed into propositions and matched against one another to determine the ways in which readers delete, reduce, or elaborate original propositions. Drum (1977) has modified the Kintsch system to five basic categories. Fredericksen (1976, 1977) has devised a system specifically for scoring inference rather than general recall data as Kintsch and Drum had done by developing a list of inferences (types and functions) that readers generate based on young children's story reading episodes.

The purpose of this study was to generate an efficacious system for scoring the output of inferences generated by adult readers. In order to create this system, two separate studies were designed to examine the inferences that readers make when they are processing texts. The first investigation was an introspective case study of two adult readers; each was asked to read two passages and to comment on the text as she/he read. The two adults were encouraged to articulate their internal observations about the texts and the mental operations they were performing while reading.* Their observations served as a basis for the creation of the scoring system that was generated.

The second study was a formal, experimental investigation of readers' recall. Passages were developed to elicit information under three sets of conditions:

1) congruence/incongruence of propositions within texts
2) wholistic/additive modes of presentation
3) free/cued recall

Method

Sample

Eight college freshmen enrolled in an English composition course participated in the study.

Preparation of Texts

A four-sentence passage containing complementary propositions was gener-

---

*Data from this ecological study have been used as referents for designing a scoring system; these data will not be presented in this study.

109

ated. The syntactic frame of this passage was used to construct a second passage that contained contradictory propositions; the concept load was controlled and matched with the first passage. Two additional passages (one complementary and one contradictory) using the original frame were also generated.

Preparation of Cards for Cued Recall

A series of 12 cards, each containing a single lexical item, was created for each passage.

Procedure

Each student was asked to read two passages in two separate presentation modes: 1) whole passage or 2) additive (sentence by sentence). Each subject received either two complementary, two contradictory, or one complementary and one contradictory passage.

After reading each passage, or each sentence in the additive presentation mode, students were asked to write recall of the text from memory. After the additive mode, subjects were asked to write a recall summary. Subjects were asked to sort the cards, containing a single lexical item, into two categories: recalled/unrecalled. When the cards were sorted, subjects were asked to reconstruct the passage by using the cards as cues for recall. The same procedure was repeated as each subject read his/her second passage.

Scoring System

The recall data from these two studies and from a previous study (Flood & Lapp, 1977) directed the generation of specific categories within the proposed scoring system. The system is based on the observed logical operations of readers. Each category and subcategory was designed from one or more bits of information recalled by our readers. This system differs from previous systems because it is based on the output of readers. Categories were only generated from readers' recall; the system was not generated from presupposed logical or psychological categories. The system is presented on the following pages.

Each category has several subcategories. The generic category and each of its subsets with explanations and exemplar are presented below:

Complete Scoring System	Example	
	Text	Recall
I. generating text identical information		
identical recall of information written in text	Jason was a lawyer.	Jason was a lawyer.
II. generating macro/micro structures creating larger or smaller units to accommodate text information		
1. synonomy-a narrowly defined category; traditionally acceptable synonyms. This category assumes a high degree of rater reliability. Synonyms can be conventionally acceptable like couch/sofa (cf. Thesaurus) or text specific	couch	sofa
2. colloquial (figurative) synonym-acceptable synonym within a specific context	policeman	cop
3. superordinate recall of the larger unit to which text element belongs	bear	animal
4. subordinate recall of small unit of which text element is a part	flower	daisy
5. categorization generation of larger concept that encompasses several text elements	uniforms, drums, batons marching people	parade
III. generating cause establishing preceding or succeeding information that can place an event within a framework that can be tolerated by the reader		

Complete Scoring System	Example Text	 Recall
1. text proactive extracting previous information from text that explains events as effects of causes	Jason was a lawyer. He became a dentist.	Jason realized he made a mistake.
2. text retroactive	Jason liked Chicago. Jason moved to Cheyenne where he enjoyed his business.	Jason moved from Chicago to Cheyenne because he didn't like his job in Chicago.
3. experience proactive presumptions about events that preceded and caused the existing event	Jason's business was successful.	Jason's family gave him a great deal of money.
4. experience retroactive assumptions about events that succeeded the existing event	Jason's business was successful.	Jason was successful as a lawyer after he sold his business and converted his business assets into client contact.
IV. generating dimension creating a spatial, temporal manner framework that can be tolerated by the reader		
1. space placing an event in space (metric or nonmetric)	Jason practiced law.	Jason's business was transcontinental, stretching from urban to rural.
2. time placing an event in time (metric or nonmetric)	Jason studied law.	In the autumn, Jason studied.
3. motion recalling movement	Jason's business was transcontinental.	Jason flew from coast to coast to help his business.
4. manner recalling specifiable characteristics	Jason studied.	Jason studied assiduously.

Complete Scoring System	Example	
	Text	Recall
V. accommodating referents establishing appropriate referents for ambiguous text elements		
1. conjunctive joining two elements	Jason was an architect. His fellow engineers praised him.	Jason was a lawyer and a dentist.
2. syncretic merging diverse elements into a single element	Jason was an architect. His fellow engineers praised him.	Jason was an architectural engineer.
3. disjunctive recall of one selected element	Jason was an architect. His fellow engineers praised him.	Jason was an engineer.
4. episodic sequencing events in a temporally fixed, irreversible order	Jason was an architect. His fellow engineers praised him.	Jason was a lawyer, but stopped being a lawyer. Then he became an engineer.
5. additive creating two sources to accommodate diverse information	Jason was an architect. His fellow engineer praised him.	One Jason was a lawyer. Another Jason was an engineer
6. anaphoric establishing a pronominal referent him.	They praised.	The engineers praised him.
VI. generating case frames creating case frameworks for text elements	Jason learned law.	Jason was taught law by the faculty of Tulane Law School.
VII. generating attributes creating modifications for actors, events, places, or dimensions		
1. actors	Jason studied law.	Mild-mannered Jason, the bookworm, studied law.
2. events attributing qualifications to events	Jason led the parade.	The parade was the grandest show in Dublin.

Complete Scoring System	Example	
	Text	Recall
3. places adding specificity to places	Jason studied in Louisana.	Jason studied in the humidity of the Southern swamps.
4. dimension attributing characteristics to dimension	He moved to Chicago.	He moved very far from Tulane.
VIII. generating text erroneous information making incorrect inferences	Jason, once a lawyer, became an accountant.	Jason studied architecture, but practiced nothing.
IX. generating text external information using established conventions of recall that conveys no new information.	Jason was a lawyer.	I don't know why Jason was what he was.

Implications

The scoring system presented in this study yielded comprehensive information about readers' recall of non-narrative passages. However, for this system to be used with narrative materials, there is a need to supplement it with a second system that may account for literary conventions that enhance or limit comprehension, e.g., knowledge of dialogue rules. Perhaps this supplementary system could be incorporated into a single system that would be useful for analyzing many different kinds of texts.

References

Drum, P. *Scoring system for recall.* Paper presented at the Second Boston University Language Development Conference, Boston, 1977.

Flood, J. & Lapp, D. *Prose analysis and the effects of staging on prose comprehension.* Paper presented at the Reading Conference of Ireland, University College, Dublin, Ireland, 1977.

Fredericksen, C. *Inference and the structure of children's discourse.* Paper presented at Society for Research in Child Development Conference, New Orleans, 1977.

Kintsch, W. On comprehending stories. In P. Carpenter & M. Just (Eds.), *Cognitive process in comprehension.* Hillsdale, NJ: Erlbaum, 1977.

EUNICE L. SCHMIDT
Seattle Pacific University

What makes reading difficult: the complexity of structures

Reading comprehension appears to be effected by the complexity of the linguistic structures used in a passage (Goodman, 1967; Ruddell, 1963; Strickland, 1962; Pearson, 1974-1975). To test this hypothesis two versions of a story, "Helen Keller" (which appeared in the *Released Exercises* of the National Assessment of Education Progress (NAEP) Reading Data, 1975) were used, a Simple Version and a Complex Version. First three methods were used to measure the complexity of the stories. Next, Cloze Versions (Simple Cloze Version/Complex Version) were designed and administered to check the complexity as a measure of reading comprehension.

Part I: Applying Three Different Methods to Measure Complexity

In the first part of this study, three different methods were used to attempt to measure the complexity of the two versions of the "Helen Keller Story". (a) Four readability formulas were used: Dale-Chall, Flesch, Fry, and Bormuth Readability. Readability formulas are frequently built on sentence length and a vocabulary measure. The formulas were used to determine whether differences in complexity between the Simple Version and the Complex Version would be reflected in readability scores. (b) Secondly, the mean number of words per T-unit was calculated. The T-unit has been defined as an independent clause with its dependent predications (Hunt, 1965). Some researchers have found the mean number of words per T-unit to be an effective measure of the complexity of writing (Hunt, 1965; O'Donnell, Griffin, and Norris, 1967; and O'Donnell, 1976). The mean number of words per T-unit increased as writing became more complex. (c) Thirdly, the Schmidt-Kittell Linguistic Complexity Scale was applied. When a study of the syntactic complexity of the spontaneous story-retell responses of 300 students to six different modes of story presentation was conducted (Schmidt, 1974), this Scale (S-KLCS) was sensitive in determining the linguistic complexity of oral and written retells. It appeared to measure within T-unit and inter-T-unit complexity more sensitively than mean number of words per T-unit. Thus is was applied to both versions of the "Helen Keller Story".

The original version of the "Helen Keller Story" as released from NAEP was used as the Simple Version in this study. Next the researcher manipulated the syntax of the story to create a Complex Version. The basic vocabulary was not changed. Prepositional phrases and participal phrases were preposed, noun structures were altered to use gerund forms; genitives were preposed; appositives and passive forms were used. For example:

PARTICIPALS:	Succeeding in arousing ... (PREPOSED)	COMPLEX
	Miss S succeeded in arousing ...	SIMPLE
GENITIVES:	Miss S spelled object's names ...	COMPLEX
	Miss S spelled the names of objects ...	SIMPLE
PASSIVES:	Speech lessons were begun ...	COMPLEX
	She began speech lessons...	SIMPLE

Applying the Readability Formulas

When the four readability formulas were applied to determine the relative complexity of each version, the grade-level scores varied from .2 to .5 grade-levels between the Simple and Complex Versions, with the Complex Version being rated slightly more complex by each scale. The designers of most readability scales call attention to the fact that the grade level obtained would be limited to being within approximately a year of the difficulty level of readers. If this statement is considered, then none of the readability scales used measured the relative complexity between the two versions of the story.

Applying the T-unit

The mean number of words per T-unit was computed to see if this might be an effective measure of the complexity of these two versions of the "Helen Keller Story". Hunt found that as students mature from grades four to twelve they use more words per T-unit. The research of O'Donnell, Griffin, and Norris supported h is data (O'Donnell, 1976). In this study, just the opposite effect appeared. The Simple Version had more words per T-unit (15.64) than did the Complex Version (13.09). T-unit length was shortened but the passage appeared to be more complex.

Applying the Schmidt-Kittel Linguistic Complexity Scale (S-KLCS)

To apply the Schmidt-Kittell Linguistic Complexity Scale, a numerical value was applied to each manipulation of the syntax. Anaphoric structures, preposing, postposing, unusual word orders, etc. were weighted. Then a syntactic complexity ratio was computed by dividing the total number or operations by the number of T-units. The Mean Ratio Score of the Simple Version of the "Helen Keller Story" was 3.68 when the Schmidt-Kittell Linguistic Complexity Scale was applied; the Mean Ratio Score of the Complex Version was 4.34. These Mean Ratio Scores were significantly different, $t(20) = 3.60$, p .01. The Schmidt-Kittell Linguistic Complexity Scale appeared to be a sensitive tool to measure the differences in complexity between these two selections.

Summary of Part I

There appeared to be no differences between versions when measured by the four readability scales. The Complex Version was measured as only .2 to .5 grade-levels more complex on any of the readability formulas.

When the mean number of words per T-unit was computed, the mean T-unit length of the Complex Version, 13.09, was shorter than the mean T-unit length of the Simple Version, 15.74. This is the opposite of frequently-cited research.

The Schmidt-Kittell Linguistic Complexity Scale appeared to be a sensitive measure of the relative complexity between the two versions. To test this sensitivity, a cloze test of comprehension was set up using the Simple and the Complex "Helen Keller Stories".

Part II: Administering the Cloze Tests of Comprehension

The cloze technique was applied to both versions of the "Helen Keller Story" so that they could be used as tests of comprehension. Directions for making and taking the tests followed Bormuth (Page, 1975). The fifth word in each passage was deleted.

The two versions (Simple Cloze Version/Complex Cloze Version) were then given to three classes of university students: two classes of senior education students enrolled in a reading/language arts course and a third class of graduate students enrolled in a reading/language arts course. A total of 84 students were tested. Students were given the Cloze Tests prior to a lecture on the use of cloze as a measure of comprehension so that they could experience the taking of such a test. Either the Simple Cloze Version or the Complex Cloze Version was distributed to the students randomly. The cloze scores were computed following Bormuth as cited. Cloze scores for each university class are given in Table 1.

Table 1

Cloze Scores of University Seniors and Graduates on the Simple Cloze Version and the Complex Cloze Version of the "Helen Keller Story"

Classes	Simple	Version Complex	Difference in Cloze Points
Summer '77 Graduates (N = 32)	66.21	48.62	17.59
Autumn '76 Seniors (N = 22)	69.50	57.80	11.70
Spring '75 (N = 31)	70.39	55.23	15.16

While both versions were well within the range of scores which would indicate that comprehension was not difficult for university students, it is interesting to note that university students would drop from 11 to 17 cloze points when given the Complex Version. (Bormuth suggested a score of 44% right could be compared to instructional level on an informal inventory, while 62% could be compared with an independent level.) A 2x3 analysis of variance for unequal N's (Winer, 1962) was applied to the data. $F(1,79) = 36.27$, p [illegible] .05, indicated highly significant differences between the two groups in cloze scores when comprehension was analyzed on the Simple and Complex Versions. No differences existed among classes, $F(2,79) = 2.2$[illegible] .05.

The Simple Version has been used in the NAEP data with fourth, eighth, and eleventh grade students. Only about half of the story was used at grade four, but is appeared to be at a relatively complex level of readability.

Overall Conclusions

The results indicated that (a) the readability scales used in this study did not appear to measure the relative complexity between the two versions of the story, (b) mean number of words per T-unit did not measure the differences in complexity between the versions either, and (c) the Schmidt-Kittell Linguistic Complexity Scale appeared to be a sensitive measure of the linguistic complexity. This was supported by the performance of university students on the cloze comprehension tests; they demonstrated difficulty in comprehension as measured by cloze scores on the Complex Version.

The researcher hypothesized that reading became more difficult because the reader processing language could not make sense out of the story; the reader could not comprehend it (Smith, 1975; Goodman, 1975). Pearson (1974-75) called for studies in which different versions of passages are constructed according to some rule-governed procedures with questions asked relevant to the structural changes affected. In this study, for one thing, the syntax was not as predictable in the Complex Version. Second, the language was less redundant (Smith, 1975). Third, more deletions occured; therefore, when the reader was processing the language there were fewer clues to enable the reader to get at the meaning. Fourth, unusual grammatical structures were used. An example of a T-unit may help to clarify these four points.

Simple Version

"*She* also wrote many books *and* articles and included an
2 4 3 5 3 2 3 3 2 1 3 1

autobiography of her early years." 15 words, S-KLCS Ratio = 3.53
6 3 3 5 3

Complex Version

"Her writing, many books *and* articles, inlcuded an
3 5 1 2 5 3 2 3 3 1 1 3 1

autobiography of early years." 12 words, S-KLCS Ratio = 4.25
6 3 1 5 3

The underlined words were those deleted in the Cloze tests. The numbers indicate the numerical weightings of the S-KLCS.

In the Complex Version, the common NV pattern was replaced with a possessive personal pronoun/gerund phrase, appositive phrase, verb. Elements of a relative clause "which were" were deleted to make the appositive phrase. The conjunctive was deleted; a comma was inserted. Other deletions included "also" and the possessive personal pronoun "her" which had identified whose "early years" were being discussed. The reader must process the gerund, "writing," acting as the subject of the verb "included" instead of the usual NV construction "she ... included". The deceivingly small changes in syntax and the use of only three fewer words appeared to make this T-unit much more complex to process.

Psycholinguistic analysis of how the reader may be processing complex linguistic sructures is of real interest. The complexity of our language makes the complex process of reading even more complex.

References

Goodman, K. S. Reading: A psycholinguistic guessing game. In H. Singer and R. Ruddell (Eds.), *Theoretical models and processes of reading*. Newark, Delaware: International Reading Association, 1970.

Goodman, K. S. What we know about reading. In P. D. Allen and D. J. Watson (Eds.), *Findings of research in miscue analysis: Classroom implications*. ERIC Clearinghouse on Reading and Communication Skills, National Institute of Education. Urbana, Ill.: National Council of Teachers of English.

Hunt, K. W. *Grammatical structures written at three grade levels*. NCTE Research Report No. 3. Champaign, Ill.: National Council of Teachers of English, 1965.

O'Donnell. A critique of some indices of syntactic maturity. *Research in the Teaching of English*, Spring, 1976.

O'Donnell, R. C., Griffin, W. J., & Norris, R. C. *Syntax of kindergarten and elementary school children: A transformational analysis.* NCTE Research Report No. 8. Champaign, Illinois: National Council of Teachers of English, 1967.

Page, W. *Helps for the reading teacher.* National Conference on Research in English Bulletin Series. ERIC Clearinghouse on Reading and Communication Skills, National Institute of Education, 1975.

Pearson, P. David. The effects of grammatical complexity on children's comprehension, recall, and conception of certain semantic relations. *Reading Research Quarterly,* 1974-1975, *10,* 155-192.

Ruddell, R. B. Language acquisition and the reading process. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading.* Newark, Del.: International Reading Association, 1970.

Schmidt, E. L. *Syntactic and semantic complexity used by children in response to six models of story presentation.* Unpublished dissertation, University of Washington, Seattle, Washington, 1974.

Smith, F. *Understanding reading: A psycholinguistic analysis of reading and learning to read.* New York: Holt, Rinehart and Winston, 1971.

Smith, F. *Comprehension and learning: A conceptual frameword for teachers.* New York: Holt, Rinehart and Winston, 1975.

Strickland, R. G. *The language of elementary school children: Its relationship to the language of reading textbooks and the quality of reading of selected children.* Bulletin of the School of Education, Indiana University. Bloomington: Indiana University, July 1962.

Winer, B. J. *Statistical principles in experimental design.* New York: McGraw-Hill Book Company, Inc., 1962.

NANCY WONG NELSON
JOHN J. GEYER
Rutgers University

Immediate memory for sentences of fast and slower readers as a function of rate of presentation

In speculations as to how mature and rapid readers read, one of the main issues is how meaning is derived from the surface representation of sentences. Huey (1968) and Smith (1971) submitted that meanings were suggested immediately by the visual forms. Athey (1971) thought that word identification processes are bypassed or telescoped in smooth reading. One step towards the understanding of such processes would be through comparative studies of fast and slower readers on what is remembered of both the meaning and surface representation of sentences. This study investigated differences between fast and slower readers in immediate memory for different aspects of sentences in connected discourse as a function of rate of presentation. The design of the study focused on the interactions between the variables 1) type of reader, 2) rate of presentation, and 3) type of test probe (meaning or non-meaning).

Studies by Huey (1968), Gates (1922), and Gilbert (1959) indicated that rapid readers are superior to slower readers in their memory for literal content. Bower (1970) found that at a fast rate of presentation of single sentences, slower readers detected changes in meaning as accurately as rapid readers. These studies, however, either did not use connected discourse or did not control the rate at which the materials were read.

Memory for both meaning and the surface representation of sentences were studied by means of a recognition paradigm developed by Sachs (1967). In studies using similar paradigms, the relationship between immediate memory for meaning and surface features of sentences differed depending on the experimental circumstances. Meaning scores were higher than (Pezdek & Roger, 1974), equal to (Sachs, 1967), or lower than (Tyler, 1971) scores measuring surface representation. The variance in results indicated that such a paradigm was serviceable in measuring differences in memory for both meaning and surface representation.

Method

Design

The study employed an ANOVA design with two levels each of type of reader (fast and slow), order of rate of presentation (eight fast passages followed by eight slow passages or eight slow passages followed by eight fast passages), amount of intervening material (0 syllables or 80 syllables), and rate of presentation (160 words per minute or 480 words per minute). These combined with four types of test probes (semantic, syntactic, lexical, identical) formed 64 possible combinations. Each subject was assigned to one of these combinations.

The 16 experimental passages were counterbalanced so that a version of each passage was given to each subject. Scores for the delayed condition were not included in the statistical analyses as these passages served merely as filler passages to discourage the readers from assuming that the last sentence of the passage was always tested. As fast and slow presentations were made in blocks, the order of rate of presentation was treated as a control variable with passages within each rate condition being randomized. The orders of presentation of rate were randomly assigned with both levels being represented an equal number of times.

Subjects

The 64 high school sophomores who served as subjects included 32 fast readers and 32 slower readers selected on the basis of Nelson-Denny Reading Test scores administered in their freshman year. The students were matched for reading comprehension. Pre-experimental analysis showed no significant differences in comprehension or vocabulary between the groups, but a significant difference in rate, $t(31) = 2.67, p < .01$. The fast group read at an average rate of 463 words per minute while the slower group read at the average rate of 246 words per minute.

Materials

The 16 passages of connected discourse used were developed by Sachs (1974) for a study on memory loss of semantic and surface features of sentences. The short form of these passages tested immediate memory whereas a form longer by 80 syllables tested delayed memory. The readability level of the passages ranged from third to eleventh grade according to Fry's readability formula (Fry, 1968) with the average at the seventh grade level. The passages ranged from 35 words to 214 words with an average length of 117 words.

Each passage contained a base sentence with that sentence presented again, changed or unchanged, as the test sentence. The test sentence, the onset of which was signalled by an asterisk, occurred immediately after the base sentence or after a delay of 80 syllables. The test sentence tested the subject's recognition memory of the base sentence, and involved a semantic, syntactic, or lexical change from the base sentence or was identical to the base sentence.

For example, if the base sentence was "The legend is that Martin once made a cloak for a poor man," the test sentences (involving respectively semantic, syntactic, lexical, and no change) would be:

"The legend is that a poor man once made a cloak for Martin."
"The legend is that Martin once made a poor man a cloak."
"The story is that Martin once made a cloak for a poor man."
"The legend is that Martin once made a cloak for a poor man."

Procedure

Students were tested individually. They were given practice passages and sets of practice sentences to assure that they would know what semantic, syntactic, and lexical changes were. Both the practice passages and experimental passages, interlaced with answer sheets, were placed before the subject, with the top passage covered by a cardboard. Two practice passages were first presented, one at 160 words per minute and the other at 480 words per minute in random order The pacing was done with a slotted cardboard with an opening large enough to expose one line at a time. Passages were structured to average eight words per line so that hand-pacing at the two rates by the experimenter was possible.

The 16 experimental passages were presented in the order and rates designated by the design of the study. After the test sentence was read, the passage was removed by the experimenter and the subject marked the answer sheet that was uncovered. The subject first indicated whether he or she thought the sentence was changed or identical and rated his or her confidence in the response on a scale from 1 to 5. This yielded a change detection score. If changed, the subject indicated the type of change: semantic, syntactic, lexical, or don't know. This yielded an identification score. Thus there were two sets of scores: change detection scores and identification scores.

Analysis

The statistical analysis consisted of separate four-way analyses of variance on the change detection scores and on the identification scores. The between group variables were type of reader and order of rate of presentation. The repeated measures were rate of presentation and probe type. Significant main or interaction effects were analyzed, when appropriate, by *t* tests on simple main effects or by the Newman-Keuls procedure for tests of significance of differences between pairs of ordered means.

Results

Change Detection Scores

No interaction effects were found. The results for the main variables were as follows:

1. No significant difference was found between fast and slower readers in their ability to detect change.

2. The mean score for the slow-fast order was significantly higher than that for the fast-slow order, $F(1,60) = 4.90, p < .05$.

3. The mean score for the slow rate was significantly higher than that for the fast rate, $F(1,60) = 10.61$, $p < .01$.

4. The main effect for probe type was significant, $F(3,180) = 2.65, p < .05$. The semantic probe was significantly higher than the syntactic, lexical, and identical probes at the .05 level. The syntactic probe was significantly higher than the lexical and identical probes at the .05 level.

Identification Scores

The results for the main variables and significant interactions were as follows:

1. Fast readers scored significantly higher than slower readers, $F(1,60) = 6.53$, $p < .05$.

2. No significant difference was found between the slow-fast and the fast-slow orders of presentation.

3. The mean score for the slow rate of presentation was significantly higher than the fast rate, $F(1,60) = 16.80, p < .01$.

4. The main effect for probe type was significant, $F(3,180) = 6.05, p < .05$. The identical probe was significantly higher than the other probes at the .05 level. (As the chances for correctly identifying an identical probe (50%) were significantly higher than the chances for correctly identifying the other probes, identification scores for the identical probe cannot be compared meaningfully with scores for the semantic, syntactic, or lexical probes).

5. The interaction between type of reader by rate of presentation was significant, $F(1,60) = 6.04, p < .05$. Whereas there was no significant difference between fast and slower readers at a slow rate, at the fast rate, fast readers outperformed slower readers, $t(60) = 3.15, p < .05$. Also, there was no significant difference between the slow and fast rates of presentation for the fast readers, but there was a significant decline in scores for the slower readers, $t(60) = 4.66, p < .05$.

6. The interaction between order of rate of presentation by rate of presentation by probe type was significant, $F(3,180) = 3.31, p < .05$. The figures drawn

depicting this interaction disclosed that the semantic probe at the fast rate at the slow-fast order behaved differently than the other probe: the score for the fast rate was higher than for the slow rate.

7. The interaction between type of reader by order of rate of presentation by rate of presentation by probe type was significant, $F(3,180)=2.66, p<.05$. An inspection of the figures drawn for this interaction followed by statistical analyses of lower order interactions disclosed that at the slow-fast order at the fast rate, the semantic probe score for fast readers increased significantly over the score at the slow rate, $t(30)=2.58, p<.05$, whereas for the slower reader at the slow-fast order at the fast rate, the identical probe score increased significantly over the score at the slow rate, $t(30)=2.08, p<.05$.

Discussion

The discussion focuses on the two main findings related to differences between fast and slower readers. The first was that slower readers can detect changes as readily as faster readers at both fast and slow rates of reading, but faster readers outperformed them in identifying changes at the fast rate. This suggests that the detection of change and the identification of change are somewhat different processes with the latter being more susceptible to disruption by excessive speed. The ability to detect change may involve a less differentiated memory of sentences while the ability to identify changes may involve a more accurate memory of particular elements in sentences. Thus, the superior ability of fast readers in process-reading material at fast rates may depend less on an undifferentiated memory of sentences than on a more accurate memory of their particular elements.

The second finding was that with a practice or warm-up period in reading, faster readers' memory for the meaning of sentences improved, whereas in similar circumstances, slower readers' memory for the surface representation of sentences improved. This accidental finding is startling in that, in both cases, the experimental groups performed better at the fast rate than at the slow rate but with respect to memory for different aspects of sentences. It suggests that readers' typical strategies and styles may not become immediately apparent but require a warm-up before they take effect, thus tending to validate warm-up practices in speed-reading courses before other exercises are undertaken. It also raises questions about the validity of speed tests that do not provide a warm-up. These results indicate that there may be some basis for the speculations that fast and slow readers differ in important ways in their processing of meaning and the surface representation of sentences.

REFERENCES

Athey, I. J. Language models and reading. *Reading Research Quarterly*, 1971, *10*, 176-181.

Bower, T. G. R. Reading by eye. In H. Levin & J. Williams (Eds.), *Basic studies on reading*. New York: Basic Books, Inc., 1970.

Fry, E. B. A readability formula that saves time. *Journal of Reading*, 1968, *11*, 513-516.

Gates, A. I. The psychology of reading and spelling with special reference to disability. *Teachers College Contributions to Education*, No. 129, 1922.

Gilbert, L.C. Speed of processing visual stimuli and its relation to reading. *Journal of Educational Psychology*, 1959, *50*, 8-14.

Huey, E. B. *The psychology and pedagogy of reading*. Cambridge: The M.I.T. Press, 1968.

Pezdek, K. & Roger, J. M. The role of comprehension in learning concrete and abstract sentences. *Journal of Verbal Learning and Verbal Behavior*, 1974, *13*, 551-558.

Sachs, J. Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics*, 1967, *2*, 437-443.

Sachs, J. Memory in reading and listening to discourse. *Memory and Cognition*. 1974, *2*, 95-100.

Smith, F. *Understanding reading: A psycholinguistic analysis of reading and learning to read.* New York: Holt, Rinehart & Winston, Inc., 1971.

Tyler, M. P. *Schizophrenics' memory for semantic and syntactic aspects of sentences occurring in connected discourse.* Unpublished doctoral dissertation, University of Texas, 1970.

PATRICIA M. CUNNINGHAM
Alamance County Schools
Graham, NC

JAMES W. CUNNINGHAM
UNC

Investigating the "print to meaning" hypothesis

One of the current "big questions" in the study of reading is what process readers go through as they read silently. Many reading researchers and theorists argue as Carver (1977-1978) does that "during rauding, most people say the words to themselves; that is, they internally articulate each successive word in a sentence" (p. 15). Others deny the existence of an intermediary speech step in fluent silent reading. Goodman (1970/1976) states that "when silent reading becomes proficient, it becomes a very different process from oral reading. It is much more rapid and not tied to encoding what is being read as speech" (p. 482). Smith (1975) states that "... reading cannot be considered a processing of 'decoding' written symbols into speech; it is neither necessary nor possible for writing to be comprehended in this way. Instead written language must be directly attacked for meaning ..." (p. 184). Goodman and Smith are not the only theoreticians who argue that reading is a print to meaning process. LaBerge and Samuels (1974/1976), while not stating that silent reading must proceed directly from print to meaning, do suggest that "we should note the possibility in the model that a visual word code may be associated directly with a semantic meaning code. That is a unit ... may activate its meaning ... without mediation through the phonological system" (p. 564). Rozin and Gleitman (1977) present a substantial amount of evidence and argument on both sides of the "print to meaning" question.

The purpose of the investigation reported here was to further empirically investigate the "print to meaning" question. In order to accomplish this purpose, the investigators designed a reading task in which using sound as an intermediary between print and meaning was most unlikely. One group of fluent readers was randomly assigned to this task. Another group was assigned to a task differing only in the fact that using sound was a possibility. The two groups were then compared on their trials to criterion in learning the passage's special lexicon, rate of reading, and recall of material read. It was reasoned that if fluent readers proceed directly from print to meaning, then removing the probability of a sound intermediary would have no effect on rate or recall.

Method

To investigate the hypothesis that fluent reading proceeds directly from print to meaning, a task was constructed which should preclude the reader's use of sound as an intermediary. Six pictures of fish were selected and names were made up for these fish. The names (*doffit, dulmet, mintex, mastib, pontud* and *pemtad*) were all six letter, two-syllable, pronounceable words. (Pronounceability was checked by asking three adult readers to say the words. Their pronunciations were immediate and consistent.) Two words were constructed for each first letter so that subjects could not discriminate on the basis of first letter only. From each of the pronounceable names, an unpronounceable name was constructed by changing the positions of the second and third letters and the fifth and sixth letters. The unpronounceable words (*dfofti, dlumte, mnitxe, msatbi, pnotdu* and *pmetda*) were verified by three adults as being unpronounceable.

Next a passage was written about these fish. The passage of approximately 450 words described imaginary properties of the six fish:

> *Pronounceable Fish Passage*
>
> There are six fish which could easily be called the "Six Wonders of the Sea World." These wonderous fish are the mintex, the doffit, the pontud, the dulment, the pemtad, and the mastib.
>
> The mintex, for example, is the fastest fish in the world. Sometimes it has been timed swimming as fast as 40 miles per hour. In a race with a mintex, any other fish would lose. Mintexes race each other sometimes as if they were sea racehorses!
>
> The doffit's claim to fame is just the opposite of the mintex. The doffit is the slowest fish in the world. It has been known to remain in the same place in the water for hours. The doffit doesn't have to worry, though, about not being able to run away from its enemy. The doffit tastes so bad that no other fish will try to eat it.
>
> The pontud is a famous fish because of the way its body lights up when it falls in love with another pontud. Scientists have tried to figure out how to make flashlights which work on the same chemical that the pontud has, but they have not succeeded. It is a beautiful sight to see two pontuds, swimming along blinking at each other!
>
> The dulmet is not an unusual fish because it is fast or slow or lights up, but because of how delicious it is to eat. Dulmet meat sells for $25 per pound in the stores where you can find it. Kings and Queens have long considered dulmet a royal dish. Dulmet is so good that some people eat it raw!
>
> The pemtad is the world's most dangerous fish because it has a poisonous bite. Pemtad poison is more deadly than rattlesnake venom and there is no antidote for it. Pemtad fishing is against the law in some countries because of its poisonous bite.
>
> Finally, there is the mastib fish. The mastib fish is so unusual because there is nothing it would rather do than be caught by a fisherman. No bait is necessary to capture a mastib. They run straight toward a hook or net. Needless to say there aren't many mastibs left!
>
> Now whether it's the speedy mintex, the slowpoke doffit, the flashy pontud, the delicious dulmet, the poisonous pemtab or the easy-to-catch mastib, you'll have to admit that here are some pretty fishy fish!

Another passage was constructed by substituting the unpronounceable fish names for the pronounceable names. Thus the first paragraph of the unpronounceable fish passage read: "There are six fish which could easily be called the 'Six Wonders of the Sea World.' These wonderous fish are the mnitxe, the dfofti, the pnotdu, the dlumte, the pmetda, and the msatbi . . ."

The study was carried out with a group of 47 fifth and sixth grade students who read at or above grade level and with a group of 14 graduate students in Reading Education. Fifth and sixth grade students were used because they were thought to be the youngest good readers who were well beyond any decoding stage which uses serial processing of words. A stratified random sample was achieved by randomly assigning the elementary and graduate students to two groups. One group learned the pronounceable words and read the pronounceable fish passage. The other group learned the unpronounceable words and read the unpronounceable fish passage.

Unfortunately, because there is a basic confounding of a word's pronounceability and its orthographic legality (Gibson & Levin, 1975), the two passages differed not only on the wanted variable of pronounceability versus unpronounceability but also on the unwanted variable of orthographic legality versus illegality. To diminish the influence of the unwanted variable without rendering the groups unequivalent, a treatment was designed which would teach each subject to associate the picture of each fish with the name as a chunk or as a sight word before the subject was asked to read the passage.

The training procedure for both groups was identical. Subjects were individually taught to associate the pictures of the fish with their names. In neither group were the names pronounced. Rather the fish were spread out, the appropriate name was placed next to each fish and subjects studied the names and the fish until they thought they could match them. Subjects continued the matching trials to criterion until they had successfully matched the fish and their names twice. Ten trials was set as the maximum number before excluding a subject from the experiment. Six of the subjects in the unpronounceable group were unable to successfully match the fish and the names after ten trials. Three subjects in the pronounceable group were unable to match them. In order to be sure that the two groups who read the passage were equivalent, three subjects in the pronounceable group who had had the greatest number of trials were excluded from the analysis.

When the subjects had successfully matched the fish and the names twice, they were given the appropriate fish passage to read. Subjects were told to read the passage so that after reading they could point to the fish that had particular traits. Subjects were timed as they read the passage and asked to read at their "normal speed."

Upon completion of the passage, the pictures of the fish were spread out and subjects were asked to point to the fish which had a particular characteristic. They were given no feedback as to the correctness or incorrectness of their responses. Next the names of the fish were spread out and subjects were asked to perform the same task with the names rather than the pictures. Finally, subjects were shown each name and asked, "What is this?" Their responses to this question were classified as *semantic* ("a fish" or "the fish whose body lights up when it falls in love" or "that big striped fish" for example) or *acoustic* (a pronunciation of any kind for the fish).

Results

The difference between the mean trials to criterion for the pronounceable group (4.75) and the unpronounceable group (5.56) was analyzed by a one-tailed t test for independent samples. This difference approached but did not reach conventional levels of significance ($p < .07$).

The differences between groups on the time required to read the passage and the two comprehension variables were measured by a two-way multivariate analysis of variance with pronounceable versus unpronounceable as one factor and age as a second factor. The multivariate F for the pronounceable versus unpronounceable factor was significant. To illuminate the reasons for this significance, univariate F's for the three dependent variables were examined. There were significant differences between the pronounceable and the unpronounceable groups on the amount of time required to read the passage, $F\ (3.43) = 6.51$, $p < .05$, and on the name comprehension variable, $F\ (3.43) = 5.65$, $p < .05$ (see Table 1). Although the picture comprehension results favored the pronounceable group, this difference was not significant.

For all analyses, there were significant differences as expected between the graduate students and the elementary students. Graduate students learned the names of the fish in significantly fewer trials, read the passages in

Table 1

Cell Means for Three Dependent

Variables across Two Factors

	Graduate Students			Elementary Students		
	Time in Seconds	Comprehension Picture	Comprehension Name	Time in Seconds	Comprehension Picture	Comprehension Name
Pronounceable	115.4	5.1	4.3	134.1	1.9	2.6
Unpronounceable	136.7	2.9	2.7	159.3	1.9	1.7

significantly less time and performed significantly better on both the picture and the name recall task. There was not a significant multivariate interaction effect.

The responses to the "What is it?" question for the elementary students were clearly affected by which names (pronounceable or unpronounceable) they learned. For the students in the pronounceable group, 13 consistently gave an acoustic response to the "What is it?" question. Only four students consistently gave a semantic response. For the unpronounceable group only six students consistently gave an acoustic response, 12 gave a semantic response. For the graduate students these results were more equivocal. Two of the seven students in the pronounceable group gave an acoustic response; three of the seven students in the unpronounceable group gave an acoustic response. This difference may be attributed to the way in which the question was put to the graduate students. The elementary students responded unhesitatingly to the "What is it?" question. The graduate students, however, were unwilling to make a response without further clarification. To clarify and to provoke a response, the investigator followed up the "What is it?" question with a "How do you have it stored?" question. Consequently, the two groups were in reality responding to two different questions.

Discussion

The rationale for this investigation stated at the beginning of this paper was that, if readers proceed directly from print to meaning without a sound intermediary, removing the probability of a sound intermediary would not result in significantly different rate and recall for the subjects who lacked this intermediary. There is evidence from the differential responses of the elementary students in the pronounceable and unpronounceable groups to the "What is it?" question that the two groups did indeed learn the pronounceable and unpronounceable stimuli differently. While the number of trials to criterion was greater for the unpronounceable group, this difference was not significant.

The results of this investigation do not support the hypothesis that readers proceed directly from print to meaning. Subjects in the pronounceable group read the passage in significantly less time than did subjects in the unpronounceable group. The magnitude of this difference is appreciated when one discovers that the elementary students in the pronounceable group took less time on the average to read the passage than did the graduate students in the unpronounceable group. Subjects in the pronounceable group also achieved significantly greater recall as measured by the name recall task.

However, there is some observational data from this study to support the print to meaning hypothesis. Subjects in both the pronounceable and unpronounceable groups appeared to enjoy the story as they were reading it. They chuckled occasionally and later asked if "there were really fish like that." In addition, the graduate students in the unpronounceable group seemed genuinely shocked when they were unable to perform well on the recall tasks. These observations lead us to hypothesize a third alternative to the "print to meaning"/"print to sound to meaning" controversy.

Readers may indeed comprehend directly from print to meaning. However, because short term memory appears to require auditory rehearsal for chunking into long term memory, recoding from meaning into sound may be necessary for recall of what was comprehended. In short, fluent readers proceed from "print to meaning to sound to memory." (We think!)

References

Carver, R. P. Toward a theory of reading comprehension and rauding. *Reading Research Quarterly*, 1977-78, *13*, 8-63.

Gibson, E. J. & Levin, H. *The psychology of reading.* Cambridge, Mass.: MIT Press, 1975.

Goodman, K. S. Behing the eye: What happens in reading. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (2nd ed.). Newark, Delaware: International Reading Association, 1976, 470-496. (Originally published, 1970.)

LaBerge, D. & Samuels, S. J. Toward a theory of automatic information processing in reading. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (2nd ed.). Newark, Delaware: International Reading Association, 1976, 548-579. (Originally published, 1974.)

Rozin, P. & Gleitman, L. R. The structure and acquisition of reading II: The reading process and the acquisition of the alphabetic principle. In A. S. Reber & D. L. Scarborough (Eds.), *Toward a psychology of reading.* Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.

Smith, F. *Comprehension and learning.* New York: Holt, Rinehart and Winston, 1975.

EDWARD FRY
JANE WEBER
Rutgers University

JOSEPH DEPIERRO
Seton Hall University

A partial validation of the kernal distance theory for readability

Most simple readability formulas take two main variables into account:

1. Vocabulary difficulty as determined by syllables or unfamiliar (low frequency) words; and
2. Syntactic complexity as determined by sentence length.

Yet, some experimentation and theory give indications that it may be possible that two sentences of identical length and identical vocabulary can have different difficulty or readability levels. For example, sentence A may appear more difficult than sentence B:

A. The fisherman, after climbing down the hill, was surprised to see the small dirty boat.

B. The fisherman was surprised to see the small dirty boat after climbing down the hill.

Hence, by taking more or other syntactic variables into account, it might be possible to develop a more accurate readability formula, at least for research purposes, if not practical application. Another possibility for practical use of such information is instructions to writers who are trying to simplify their written material.

Several studies such as those done by Coleman (1964; 1965) indicate that "embedded" sentences such as Example A where the subject and verb are split by "distance" in form of a phrase are harder to read than non-embedded sentences. A linguistic theory proposed by Yngve (1968) suggested that "distance" (a word or phrase) coming before the subject and verb would make the sentence more difficult than the same distance coming after the subject and verb. Using this information and a small scale pilot study, the senior author proposed the Kernal Distance Theory (Fry, 1975) which stated:

> ". . . that the nearer the kernel is to the beginning of the sentence, the easier the sentence, and the less distance between elements of the kernel, the easier the sentence. For the purposes of this theory, we are calling the kernel the subject element, the verb element, and when present, the object element. . . There is a minor and further refinement of the Kernel Distance Theory which states that the distance between the verb and object causes less difficulty than the distance between the subject and verb. Distance can be defined as number of words, though in practice it is often the embedding of a phrase or clause."

It might be noted that the term "distance" is somewhat unique to this study and is defined as a word or phrase. The term "kernel" refers to the subject, verb, and sometimes object of a sentence and may not carry the same interpretation as used by Chomsky (1957; 1965).

Method and Results of Two Studies

This paper will report the results of two different investigators to verify the

Theory on four different populations using four different measures as dependent variables.

First Study

Method. DePierro (1976) developed 16 pairs of sentences, four for each part of the theory (see Table 1) and asked 30 college students and 30 elementary pupils to individually read them. He measured:

1. Ability to recall the exact *words* in the sentence;
2. The *time* it took to read silently the complete sentence;
3. The *delay* between the end of the silent reading of the sentence and the onset of oral recall.

Sentence pairs were written so that each pair contained the same words and only syntax was altered according to hypotheses. A balanced rotated design was used to cancel out order effects.

The elementary pupils had reading abilities between 5.0 and 7.0; there were no remedial readers. The 14 male and 16 female college subjects scored between 7 percentile and 94 percentile on the Nelson Denny Reading Test, and hence represented the full range of students found in college settings. The test of significance was the Wilcoxon *t*.

Results. DePierro found that splitting the subject and verb (D_2) often significantly increased sentence difficulty. See results for $H^{O}2$ and $H^{O}3$ in Table 1. However, he found no support for a difference between the subject-verb split and the verb-object split. With the college population only he found a difference contrary to the theory that distance after the kernel was more difficult than distance before.

Second Study

Method. In the second study, Weber (1977) using good and poor readers in a junior college, attempted to replicate DePierro's findings using a different dependent variable. She used student judgment of sentence pairs as to which of the two sentences was most difficult. A group test was constructed using DePierro's 16 pairs of sentences and another 16 pairs of sentences of her own to make the group test more reliable by increasing its length.

The 52 poor readers were students who scored below 43 on the Comparative Guidance and Placement Program battery of tests developed by Educational Testing Service and were placed in three Basic Composition classes. The 53 good readers scored above 43 on the same test and were placed in three Language and Literature classes. The test of significance was the Chi Square.

Results. Weber basically confirmed DePierro's findings; distance between the subject and verb was judged more difficult than the same distance either before or after the kernel. No support was found for a difference between distance before or after the kernel. Nor was there a difference for distance between the subject and verb, and the distance between the verb and the object (See Table 1).

Discussion

The Kernel Distance Theory, as tested by these studies, can now be simplified to six words, DON'T SPLIT THE SUBJECT AND VERB. Stated a bit more technically, distance between the subject and verb significantly increases the difficulty of a sentence over one with the same amount of distance occurring either before or after the kernel. These differences showed up on

Table 1

Summary of Agreement With Hypotheses of the Kernel Distance Theory

Fry Hypothesis	Weber[a] 32 Sentence Pairs				DePierro[b] 16 Sentence Pairs	
	College				Elementary	College
	Poor N=52	Good N=53	Total N=105		N=30	N=30
H_o^1 $D_1 > D_4$	+	-	-	Words	-	-.01
				Time	+	-.05
				Delay	-	+
H_o^2 $D_2 > D_4$	+	+	+.05	Words	+.05	-
				Time	+.01	+.05
				Delay	+	+
H_o^3 $D_2 > D_1$	+.05	+.05	+.05	Words	+.05	-
				Time	+.01	+.05
				Delay	+	+
H_o^4 $D_2 > D_3$	+	-	-	Words	-	-
				Time	+	-
				Delay	-	-

[a] Plus sign (+) means agreement with hypothesis using judgment, sig. where indicated by X^2, p value.

[b] Plus sign (+) means agreement with hypothesis on three measures, sig. where indicated by Wilcoxon t, p value.

[c] Subject Verb Object
D_1 D_2 D_3 D_4

four different measures of difficulty with four different populations (though not entirely consistently).

This finding is in agreement with the embedded sentence study of Coleman (1964) who had college students memorize pairs of sentences from a memory drum. It also agrees with the Hamilton and Deese (1971) listening comprehension study of center-embedded sentences. And it is in accord with the syntactic measures of Marcus (1971) and Botel and Granowsky (1972).

The other parts of the Kernel Distance Theory are not confirmed nor does the data reported here confirm Yngve's theory concerning the memory burden difficulty supposedly caused by placing a word or phrase at the beginning of a sentence.

However, there was some tendency for all hypotheses to be seen as positive. Weber's poor readers all performed in the direction hypothesized and DePierro's elementary pupils on the time measure tended to support the hypotheses. If any future research is done on these hypotheses, it is recommended that it be done with poor readers.

It is doubtful that the finding of these studies will modify the practical or more popular readability formulas, but it might have some bearing on future readability research.

These studies do have one very practical outcome. Writers should avoid subject-verb splits if they wish to simplify their writing.

> ". . . that the nearer the kernel is to the beginning of the sentence, the easier the sentence, and the less distance between elements of the kernel, the easier the sentence. For the purposes of this theory, we are calling the kernel the subject element, the verb element, and when present, the object element . . . There is a minor and further refinement of the Kernel Distance Theory which states that the distance between the verb and object causes less difficulty than the distance between the subject and verb. Distance can be defined as number of words, though in practice it is often the embedding of a phrase or clause."

Results. DePierro found that splitting the subject and verb (D^2) often significantly increased sentence difficulty. See results for $H_0 2$ and $H_0 3$ in Table 1.

References

Botel, M. & Granowsky, A. Formula for measuring syntactic complexity. *Elementary English*, 1972, *49*, 513-516.

Chomsky, N. *Syntactic Structures*. The Hague: Mouton, 1957.

Chomsky, N. *Aspects of the Theory of Syntax*. Cambridge, Mass.: M.I.T. Press, 1965.

Coleman, E. B. The comprehensibility of several grammatical transformations. *Journal of Applied Psychology*, 1964, *48*, 186-190.

Coleman, E. B. Learning of prose written in four grammatical transformations. *Journal of Applied Psychology*, 1965, *49*, 332-341.

DePierro, J. P. *Some effects of sentence structure variables on reading ease*. Unpublished doctoral dissertation, Rutgers University, 1976.

Fry, E. B. A kernel distance theory for readability. In G.H. McNinch and W.D. Miller (Eds.), *Reading: Convention and inquiry*. Twenty-Fourth Yearbook of the National Reading Conference. Clemson, South Carolina: The National Reading Conference, 1975.

Hamilton, H. W., & Deese, J. Comprehensibility and subject verb relations in complex sentences. *Journal of Verbal Learning and Verbal Behavior*, 1971, *10*, 163-170.

Marcus, A. D. The development of a diagnostic test of syntactic meaning clues in reading. In R. E. Leibert (Ed.), *Diagnostic viewpoints in reading*. Newark, Delaware: International Reading Association, 1971, 48-63.

Weber, Jane. *The kernel distance theory – Evaluation by student judgment of sentence difficulty*. Unpublished master's thesis, Rutgers University, 1977.

Yngve, V. H. A model and a hypothesis for language structure. *Proceedings of the American Philosophical Society*, 1960, *104*, 444-466.

LAURA CHODOS
The University of the State of New York

PETER MOSENTHAL
S.U.N.Y. at Albany

Fourth graders' comprehension of story structures under three recall conditions

Text, or connected discourse, can be viewed at two levels: (1) the microstructure level and (2) the macrostructure level (cf. van Dijk, 1977). At the microstructure level, one is concerned with understanding how lexical, propositional, and sentential variables influence text comprehension. At the macrostructure level, one is concerned with understanding how sets of propositions and sentences are structured relative to one another and how this relative structuring influences text comprehension.

For the most part, reading research has focused on the microstructure level, investigating word recognition variables and sentence comprehension variables, and has paid little attention to the macrostructure level. Yet, as studies in other nonreading disciplines have demonstrated, it has become increasingly apparent that models of word recognition or sentence comprehension cannot account for many of the factors influencing children's comprehension of written text. Although theories of written text comprehension must eventually explain how these macrostructure units influence children's reading comprehension of a text's macrostructure, the significance of macrostructure variables must be clearly understood before this relationship can be explored.

Researchers investigating how people comprehend story macrostructures (e.g., Glenn, 1978; Mandler & Johnson 1977; Stein, 1978; Stein & Glenn, 1978; Stein & Nezworski, 1978) have assumed that there exists an "ideal" story macrostructure which can be described in terms of a specified ordering of story categories, logically and temporally related. For example, Mandler and Johnson (1977) have defined the ideal story macrostructure as being comprised of four categories: (1) Setting, (2) Beginning, (3) Development, and (4) Ending. The Setting is a group of propositions or sentences which accomplish the following: They introduce the main characters of the story, they describe the time and local of the story, and they provide additional information the reader needs to know to understand the events that follow. In the Beginning, something causes the main character to respond in some way (This response often results in some type of displacement from the character's normal routine). The Development describes the actions a character uses to obtain some explicitly or implicitly defined goal. Finally, the Ending describes the consequences of the effect of the goal either having been achieved or not achieved.

A second assumption of story grammarians is that these macrostructures have a real-time psychological counterpart existing in mind; these grammarians have called these mind macrostructures "schemata." In sum, story schemata consist of sets of expectations about stories, about the categories of which they are composed, the way in which these categories are sequenced, and the types of connections between categories that are likely to occur (cf. Mandler, 1978; Stein, 1978; Stein & Glenn, 1978).

The question of how children employ schemata to comprehend written text is unresolved for several reasons. First, the early work of Piaget (1926/1960) and Fraisse (1963) suggests that young children lack a framework for organizing the events of a story. Both Piaget and Fraisse reported temporal inversions in young children's recall and assumed the inversions resulted from an inability to establish a framework for organizing chronological, casual, and deductive relations. Such findings suggest that children do not have the benefit of well defined schemata for organizing events.

In contrast to these findings, other findings (e.g. Brown, 1975, 1976; Stein, 1978; Stein & Glenn, 1978) have noted that children do, in fact, organize narrative information according to well defined schemata. For example, in one study, Brown (1975) tested kindergarten and second-grade children's recognition, reconstruction, and recall of simple narrative sequences. For all children, logical sequences were better retained than were arbitrary sequences. The difference between the two groups was that although kindergarten children could recognize and reconstruct the correct sequences, they had difficulty in maintaining the correct order during recall. Second-grade children, on the other hand, were able to establish the correct order of the events no matter what the method of testing was.

In addition to the problem of such contradictory findings, two other problems with studies investigating children's use of schemata in comprehending stories stem from the conditions and procedures used to operationally define story comprehension. In terms of conditions, most research has focused on children's comprehension of orally presented stories. The question of how children comprehend written stories remains largely unexplained. Although many researchers have adopted the assumption that processes underlying listening comprehension are the same as those underlying reading comprehension (e.g., Goodman & Goodman, 1977), other researchers have demonstrated that this assumption is neither tenable at the microscopic nor the macroscopic level of processing (e.g., Hildyard & Olson, 1978; Mosenthal, 1978; Thorndyke, 1977).

In terms of procedures, the most typical way story comprehension is defined is by having subjects recall a story immediately after presentation. Two measures of organization in recall are usually derived from this operational procedure. The first is the relative salience of categories. As shown by Glenn (1978), Stein and Glenn (1978), and Mandler and Johnson (1977), categories vary in terms of their salience or structural importance in recall. In particular, Beginning and Development categories tend to be recalled more often than Setting and Ending. A second measure of organization focuses on the way the categories are grouped in recall. According to story grammars, events should be recalled in the sequence of Setting, Beginning, Development, and Ending.

Studies, such as Brown's (1975), have demonstrated that under free recall conditions, young children tend to use a schema organization for remembering and ordering events. On the other hand, under such conditions as recognition and forced recall, children tend to episodically remember information without the benefit of organizing schemata. In short, whether children would use a schemata to order episodes from a story in operational procedures other than free recall in unclear.

Another point which needs investigation is whether children's schemata for a story changes over time. Because most studies of children's comprehension of stories focuses on immediate recall, it is unclear whether this schema structure prevails over time. Research is divided on this point. On the one hand, there are those researchers (e.g., Mandler, 1978; Spiro, 1977) who argue that over time, children will reconstruct narrative sequences so that they are assimilated to new text structures, based upon redefined schemata. On the other hand, there are those researchers (Cofer, 1943; Rubin, 1977; Squire, Chase, & Slater, 1972) who argue that the structure of narrative sequences is left in tact due to superficial, or reproductive processing. In sum, how pervasive children's story schemata are in organizing related narratives into a larger schema structure has been poorly addressed.

The purpose of this study was to investigate the function which schemata play in fourth grade children's reading comprehension of stories under different comprehension conditions. The three different comprehension conditions included: immediate recall, delayed recall, and immediate forced recall. If children do, in fact, use schemata to comprehend stories, as story grammars predict, then one would expect children, in an immediate recall condition, to consistently recall more information from some story categories than from others. Secondly, if schemata are the principal means by which children organize and interpret narrative events, one would predict that the most saliently remembered categories in free recall would be the most saliently remembered categories in the forced recall. Finally, one would predict that if children's schemata serve to direct reconstructive, long-term remembering, then four complete stories which, when put together form a complete new story, will be remembered in the order of saliency of the categories they represent. In other words, if four stories, when put together, comprise four categories of a well formed new story (these categories being, namely Setting, Beginning, Development and Ending), then one would predict that in a delayed recall of these stories, the stories most saliently recalled would be those which comprise categories most saliently recalled in the individual story recall conditions. These hypotheses were tested in the following experiment.

Method

Subjects

Thirty students were randomly selected from two fourth-grade classrooms in a large suburban school near Albany, New York. All students participating in the experiment could read and respond to the stimuli in the manner required.

Materials

A story from a fourth-grade social studies text was rewritten into four smaller stories following four different episodic parts of the original story. Each of the four stories was written to conform to the story-grammar tree structures of Mandler and Johnson (1977). Each story structure consisted of four categories: the Setting, the Beginning, the Development, and the Ending. There were three events in the Setting, three in the Beginning, four in the Development, and one in the Ending. Each event within and across all four stories consisted of eleven propositions, as defined by Kintsch (1974). The words in the propositions within and across all four stories were controlled for in terms of frequency (Kučera & Francis, 1967) and imagery value (Paivio, Yuille, & Madigan, 1968).

Procedure

The thirty students read each of the four stories. After reading a story, each student was first required to write all that he could remember about the story. After the written recall, each student attempted to complete a cued recall test. This test required that the student fill in deleted information in eleven statements. Each of the statements corresponded to information in each of the eleven events in a story, respectively. All students were told that they should read the story carefully, since they would be asked to write about each story and answer several questions. In addition, all students were informed that they would be asked to write about the story one week later. The students read and completed the tasks for each of the four stories on consecutive days of one week. Eight days later, all students were asked to recall as much as they could about all four stories. The students were tested in groups, with each group receiving similar instructions. Instructions for the free recall were, "Write everything you can remember about the story you have just read. Try to remember everything, including the words the author used." For the delayed recall, subjects were instructed to "Write everything you can remember about the four stories you read last week. Try to remember everything, including the words that were in the story."

Scoring

Responses to the free and cued recall tasks were marked by two raters. Acceptable responses included propositions which were literal reproductions from the text stimuli or else were synonymous with predefined text propositions. In the immediate free recall, four scores were tallied for each story. These scores included the number of propositions recalled in the Setting, the Beginning, the Development, and the Ending. Similarly, in the immediate cued recall, the number of propositions correctly recalled from the Setting, Beginning, Development, and Ending were likeways tallied. In the delayed recall, the number of propositions correctly recalled from each of the four categories from each of the four stories was also tallied.

Results

In the following analyses, in order to compensate for the fact that some story categories had more events and propositions associated with them, recall and recognition were calculated in terms of the mean proportion of propositions recalled and recognized per category. The mean proportion of propositions recalled per category was calculated by multiplying the number of events per category times the number of propositions per event and dividing this product into the number of propositions recalled per category. In determining mean proportion of propositions recalled per category, the number of cued recall questions correctly answered was divided by the number of events per category.

The first hypothesis was that if children use schemata to comprehend stories, then children should consistently recall more propositions from some story categories than from others. An Analysis of Variance of the mean proportion of propositions in basic story categories recalled immediately after reading the text (see Table 1) supported this hypothesis, $F(3, 116) = 19.11$, $p < .01$. No interactions proved significant. This hypothesis was further supported by a series of Scheffé pairwise comparisons between the Setting and the other categories within the stories; all Scheffés were significant except for the pairwise comparison between the Setting and the Beginning in Story 2. Propositions from the Setting and Beginning were recalled significantly more frequently than from the Development; propositions from the Development were recalled significantly more frequently than from the Ending across all four stories, with the exception noted in Story 2.

The second hypothesis posited that if schemata are the principal means by which children organize and interpret narrative events, then the most saliently remembered categories in free recall should be the most saliently remembered categories in the forced recall. This hypothesis was supported by the fact that an overall Analysis of Variance of the mean proportion of cued recall statements correctly answered was significant, $F(3, 116) = 17.7$, $p < .01$ (see Table 1). No interactions proved significant. A series of Scheffes proved significant for all pairwise comparisons, except between the Setting and the Beginning in Story 2 and between the Beginning and the Development in Stories 3 and 4. The fact

Table 1

Saliency of Category and Story Recall over Four Stories Recalled under Three Different Conditions

Saliency (measured as mean proportion of propositions correctly recalled) of Story Categories in Immediate Recall

	Setting	Beginning	Development	Ending
Story 1	.72	.51	.43	.22
Story 2	.67	.58	.48	.17
Story 3	.61	.49	.39	.16
Story 4	.69	.53	.42	.20

Saliency (measured as mean proportion of recall questions correctly answered) of Story Categories in Forced Recall

	Setting	Beginning	Development	Ending
Story 1	.69	.58	.24	.20
Story 2	.58	.53	.39	.12
Story 3	.59	.42	.37	.19
Story 4	.63	.47	.39	.23

Saliency (measured as mean proportion of propositions correctly recalled) of Story Categories in Delayed Recall

	Setting	Beginning	Development	Ending
Story 1	.42	.21	.29	.34
Story 2	.31	.10	.18	.21
Story 3	.17	.07	.12	.18
Story 4	.05	.02	.10	.12

that the saliency of the categories in the immediate and delayed recall conditions was similar supports the contention that children by fourth grade use the same schemata for comprehending stories under varying operational conditions.

The third hypothesis was that if children's schemata serve to direct reconstructive, long-term remembering, then children should recall four stories, serving as a Setting, Beginning, Development, and Ending, respectively, to a superordinate story, in the same manner as they recalled the Setting, Beginning, Development, and Ending in the immediate recall stories. An Analysis of Variance of the number of propositions recalled from each story revealed an overall significance between stories, $F(3, 116) = 21.82$, $p, < .01$ (see Table 1). A series of Scheffé pairwise comparisons revealed that Story 1 (i.e. the Setting) was recalled significantly better than Story 2 (i.e. the Beginning); Story 2 was recalled significantly better than Story 3 (i.e. the Development); and Story 3 was recalled significantly better than Story 4 (i.e. the Ending). These results were similar to those of the immediate recall findings that demonstrated a saliency order, from most to least, of Setting and Beginning, Development, and Ending. A series of Scheffes revealed a significantly different saliency preference for recalling categories within each story in the delayed recall condition. The order of saliency preference for each of the stories was Setting, Ending, Development, and Beginning.

The uniqueness of this final finding is further compounded when one compares the recall order of categories within the stories in the immediate and delayed recall conditions and between stories in the delayed recall condition. In the immediate recall condition, the Setting was recalled first 96% of the time, the Beginning was recalled second 97% of the time, the Development was recalled third 94% of the time, and the Ending was recalled fourth 99% of the time. In the delayed recall, between stories, Story 1, as Setting, was recalled first 97% of the time; Story 2, as the Beginning, was recalled second 98% of the time; Story 3, as the Development, was recalled third 94% of the time; and Story 4, as the Ending, was recalled fourth 99% of the time. The recall order of these categories was in keeping with the order story grammars predicted. However, in the delayed recall within stories, the Setting was recalled first 94% of the time, the Ending was recalled second 86% of the time, the Development was recalled third 95% of the time, and the Beginning was recalled fourth 92% of the time. (A series of ANOVAs performed on the recall percentage of categories under each condition all proved significant, $p < .01$) This order of category recall was not in keeping with the order story grammars predicted.

Discussion

The findings of this study support the notion that children employ schemata in comprehending written stories much the same as they do in comprehending spoken stories in immediate recall. The nature of these schemata is such that text is processed hierarchically, with the most significant story categories being the best recalled (cf. McKoon, 1977; Meyer & McConkie, 1973). Furthermore, it appears that schemata serve to temporally order categories recalled, with the most salient categories being recalled first.

These findings additionally demonstrate that fourth-grade children employ schemata in the same manner under both free and forced immediate recall conditions.

While these conclusions are supported by the findings in the immediate recall condition, they are not wholly supported by the findings in the delayed condition. In the delayed condition, individual stories serving as episodes of a larger story were recalled in the manner story grammars predicted; saliency and order of recall followed the predicted pattern. On the other hand, saliency and order of recall of categories of the individual stories did not conform to the predicted pattern. The order of saliency and recall was Setting, Ending, Development, and Beginning.

This deviation in the delayed condition is unique for several reasons. First, although the curve representing the saliency of recall resembles the typical serial learning curve, with recency and primacy effects, the recall order of the Setting and Ending do not conform to the typical recall order in serial position studies. In addition to the recall order finding, other studies (e.g., Johnson & Scheidt, 1977; Myer & McConkie, 1973) have demonstrated that primacy effects in prose recall is not a function of seriation. What appears to have taken place is that, over time, fourth graders assimilate smaller story schemata into larger schemata and, in the process, reorganize the original schematic organization of individual story episodes. However, why this study's children arrived at the particular category organization they did in the delayed recall needs further investigation and explanation.

These findings raise an interesting question for the teaching and evaluation of reading. Most teaching and evaluation is based upon teaching and evaluating children's immediate comprehension of text. Comprehension, however, is not only a process of a moment but of a life time. The question of how one teaches children to comprehend text episodes so that they are effectively and "ideally" integrated over time needs careful consideration. How one is to evaluate the effectiveness of this integration also needs to be considered.

References

Brown, A. Recognition, reconstruction, and recall of narrative sequences by preoperational children. *Child Development*, 1975, *46*, 156-166..

Brown, A. Semantic integration in children's reçonstruction of narrative sequences. *Cognitive Psychology*, 1976, *8*, 247-262.

Cofer, C. N. A comparison of logical and verbatim learning of prose passages of different lengths. *American Journal of Psychology*, 1941, *54*, 1-20.

Fraisse, P. *The psychology of time*. New York: Harper and Row, 1963.

Gardiner, J. M., Passmore, C., Herriot, P. & Klee, H. Memory for remembered events: Effects of response mode and response-produced feedback. *Journal of Verbal Learning and Verbal Behavior*, 1977, *16*, 45-54.

Glenn, C. G. The role of episodic structure and story length in children's recall of simple stories. *Journal of Verbal Learning and Verbal Behavior*, 1978, *17*, 229-247.

Goodman, K. S. & Goodman, Y. M. Learning about psycholinguistic processes by analyzing oral reading. *Harvard Educational Review*, 1977, *47*, 317-333.

Hildyard, A. & Olson, D. R. Memory and inference in the comprehension of oral and written discourse. *Discourse Processes*, 1978, *1*, 91-117.

Johnson, R. E. & Scheidt, B. J. Organizational encodings in the serial learning of prose. *Journal of Verbal Learning and Verbal Behavior*, 1977, *16*, 575-588.

Kintsch, W. *The representation of meaning in memory*. Hillsdale, N. J.: Erlbaum, 1974.

Kucera, H. & Francis, W. N. *Computational analysis of present-day American English*. Providence,

Squire, L. C., Chace, P. M. & Slater, P. C. Assessment of memory for remote events. *Psychological*

McKoon, G. Organization of information in text memory. *Journal of Verbal Learning and Verbal Behavior*, 1977, *16*, 247-260.

Mandler, J. M. A code in the node: The use of a story schema in retrieval. *Discourse Processes*, 1978, *1*, 14-35.

Mandler, J. M. & Johnson, N. S. Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 1977, *9*, 111-151.

Meyer, B. J. F. & McConkie, G. W. What is recalled after hearing a passage? *Journal of Educational Psychology*, 1973, *65*, 109-117.

Rubin, D. C. Very long-term memory for prose and verse. *Journal of Verbal Learning and Verbal Behavior*, 1977, *16*, 611-621.

Spiro, R. J. Remembering information from text: The "state of schema" approach. In R. C. Anderson, R. J. Spiro & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge*. Hillsdale, N. J.: Erlbaum, 1977.

Squire, L.C., Chace, P. M. & Slater P. C. Assessment of memory for remote events. *Psychological Reports*, 1975, *37*, 223-234.

Stein, N. L. The development of appreciation of stories in children. In S. Madeja (Ed.), *The arts, basic skills, and cognition* (Vol. 2). St. Louis, Mo.: CEMREL, 1978.

Stein, N. L. & Glenn, C. G. An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.) *Advances in discourse processes*, Vol. 2, *New directions*. Norwood, N. J.: Ablex, 1978.

Stein, N. L. & Nezworski, T. The effects of organization and instructional set on story memory. *Discourse Processes*, 1978, *1*, 177-193.

Thorndyke, P. W. Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 1977, *9*, 77-110.

van Dijk, T. A. Semantic macro-structures and knowledge frames in discourse comprehension. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, N. J.: Erlbaum, 1977.

MARGARET SMITH-BURKE
PATRICIA S. GINGRICH
DANIEL EAGLEEYE
New York University

Differential effects of prior context, style and deletion pattern on cloze comprehension

Psycholinguists and information processing theorists (Goodman, 1976; Rummelhart, 1976; Smith, 1975) suggest that the proficient reader who possesses an adequate knowledge base relies on context and syntactic and semantic cues in order to comprehend reading material. Context allows the reader to access relevant knowledge which in turn, facilitates the comprehension process (Schallert, 1975). It has been demonstrated that interpretation of entire passages can be affected by the context (e.g., the passage title, Bransford and Johnson, 1972; Schallert, 1975).

Goodman (1976) proposes that, as readers progress through a passage, they must utilize their own semantic base to recreate the author's message. Implicit in this model of the reading process is the idea of conceptual build-up. As readers continue reading, the availability of additional information, cued from the text, further constrains the possible interpretations of subsequent text. In analyzing the miscues of fourth grade children reading narrative material, Menosky (1971) found that the quality of the students' miscues changed as they read the first, second, and final thirds of the story. Miscues progressively detracted less from the author's message. Menosky interpreted her results as evidence of contextual build-up in support of the Goodman model.

If the cloze procedure (Taylor, 1953, 1956) as a measure of comprehension has validity, according to the psycholinguistic models, evidence of contextual build-up should be expected. Some studies suggest that the ability to complete cloze is bilaterally constrained by four to five words on either side of the deletion (Aborn, Rubenstein, & Sterling, 1959; MacGinitie, 1961; Miller & Coleman, 1967). Other studies provide evidence that this ability may be influenced by context that extends beyond the immediate sentence, perhaps even to an entire paragraph (Ramanauskas, 1972). No consensus has been reached concerning the effect of context on cloze nor has this influence been studied over passages longer than a paragraph.

The purpose of this study was to examine the effect of varying amounts of prior context on cloze performance. Consistent with the Goodman model, it was predicted that cloze comprehension would be greatest for those who received the most context prior to the cloze passage (500 words), less for those who received some context (250 words), and least for the control group, if contextual build-up occurred.

Two other factors, style of writing (narrative or expository) and deletion pattern (one-fifth or function word), were considered. Sager (1977) suggests that in technical material the logical argument is carried by function words. Due to the descriptive nature of narrative material, the content usually is not organized to present a logical argument and function words serve a more general purpose. Therefore, it was hypothesized that function words would be more difficult to replace in expository material than in narrative material. One-fifth deletion was included so that the results of this study could be compared with the existing research on cloze.

Method

Sample

The 93 subjects for this study consisted of 53 eleventh-grade students and 40 twelfth-grade students in a suburban New Jersey high school. They were enrolled in either academic composition or general English classes. The sample mean on a standardized reading comprehension test ranged from the 38th to 40th percentile.

Materials

Two narrative cloze passages (*Kon-Tiki* and *1984*) and two expository cloze passages (*How to Read Body Language* and *The Good-Natured Gorilla*) were created. On three readability measures, the narrative material ranged from seventh to ninth grade level and the expository from ninth to twelfth grade. A one-fifth deletion pattern was applied to *Kon-Tiki* and *Body Language*. For *1984* and *Gorillas*, if the fifth word (or multiple of 5) was a function word, it was deleted. If not, the next function word following the fifth word was deleted. All cloze passages were approximately 250 words long allowing for 50 deletions. All deleted words were replaced by un-numbered 10-space blanks.

Each of these four passages appeared in three prior context conditions: (a) only the cloze passage, (b) the cloze passage preceded by a page of 250 words of the text, and (c) the cloze passage preceded by two pages of the text, 250 words per page.

Procedure

This study was conducted in six 45-minute sessions over a period of three to four weeks. During the first session two teaching passages were administered and four points were stressed: 1) Only one word belongs in the blank. 2) Spelling does not count. 3) If more than one word seems to fit the blank space, choose the one you think is most consistent with the author's message and style. 4) Remember that cloze is a difficult task. These four points were also reviewed at the beginning of each cloze testing session. In the second session the subjects were administered the Reading Comprehension section of the *Iowa Silent Reading Test*, Level 3, Form F (1972). In each of the last four sessions, subjects completed the experimental cloze passages.

Design

A split-plot design was used with amount of prior context the between-group variable. Style and deletion pattern were repeated within-group variables counterbalanced for passage order, grade and class.

Passages were scored for percent of exact replacement of deleted items. All two-word responses, morphological variations, and omissions were scored incorrect.

Results and Discussion

Preliminary investigations revealed no differences attributable to sex or grade. Therefore, these variables were not considered further.

In a 3 x 2 x 2 (Context x Style x Deletion) factorial analysis of variance, no significant main effects of prior context were found, $F(2,90) = 0.365, p > .05$. This result seems to lend support to the position that the cloze task forces the subject to concentrate on a more limited area of the text (Aborn et al., 1959; MacGinitie, 1961; Miller & Coleman, 1967). Another interpretation of these data from the perspective of discourse analysis is that the demands of the cloze task may be interfering with the subject's ability to utilize the strong cuing systems of connected discourse, forcing the subject to attend to syntactic and semantic information at the sentence level. If this interpretation is accurate, it appears that cloze is not an adequate measure of the reading process.[1]

Although there was no overall effect, upon inspection of individual items, it appeared that certain types of items such as pronouns and adjectives of quantity were more affected by prior context. One way to identify these items would be to analyze cloze answers using the lexical marker-transfer feature model proposed by Finn (1977-78). Once words of this type are identified, meaning those which are affected by prior context, experimental passages could then be constructed using these words as deleted items, thus modifying the cloze deletion pattern as it has been traditionally defined.

Style, F (1,90 = 60.56, $p < .001$, and type of deletion, F (1,90) = 1873.67, $p < .001$, were both significant. However, even though the difference in style was significant, it only accounted for 3% of the total sums of squares. The score for narrative style (M = 54.75) was significantly greater than that for expository style (M = 48.43). Although readability was equivalent within style, it was not equivalent across styles. Therefore, this difference in means may be a function of the different readability ranges or the style.

The significant difference for type of deletion accounted for almost 53% of the total sums of squares. Table 1 presents the means. In both styles, one-fifth cloze was more difficult than function word cloze.

Table 1

Mean Percent of Correct Responses

on Narrative and Expository Passages

with Random One-Fifth or Function Word Deletions

	CTX 1		CTX 2		CTX 3		TOTAL *M*	
	1/5	Fct	1/5	Fct	1/5	Fct	1/5	Fct
Nar	39.48	66.77	42.26	68.71	41.42	69.87	41.05	68.45
Exp	35.94	59.94	36.77	63.68	33.68	60.58	35.46	61.40

It was hypothesized that there would be a significant interaction of Style x Deletion, that subjects would score higher on function word deletions in the expository style passage than in the narrative style passage. Results show no significant interaction, F (1,90) = 1.026, $p > .05$. In this study the function word deletion pattern was constructed to be as consistent as possible with the cloze literature. Inspection of the deleted items revealed a low proportion of words such as conjunctions which Sager (1977) notes primarily carry the logical argument of the text. Further research employing a deletion pattern based only on conjunctions might produce evidence of the specific role of these functions words in expository material.

---

[1] The authors wish to thank reviewer 8.1 for helpful comments on this point.

A major characteristic of narrative style is the elaboration — the redundancy, the variety of vocabulary items, the use of synonyms, the reliance on a number of terms to describe a particular concept. For example, in *Kon-Tiki, wave* is referred to as (a) "the wall," (b) "the great green wall," (c) "the mountain," (d) "the hell," and (e) "the sea." A characteristic of expository material, on the other hand, is precision. Lexical items, particularly nouns, are not varied in an attempt to maintain clarity.

A further analysis (based on an analysis by Fillenbaum, Jones & Rapoport, 1963) of the one-fifth cloze passages investigated this style difference. A lexical score, representing the percent of lexical items correct, and a function score, representing the percent of function words correct, were calculated (see Table 2). A 3 x 2 analysis of variance (Context x Style) was performed using each of these scores, lexical and function percentages. Results indicate that lexical items in narrative material were significantly more difficult to replace than in expository material, $F(1,90) = 5.40$, $p < .05$. This further analysis also indicated that function words in the expository material were significantly more difficult to replace than those in narrative material, $F(1,90) = 6.56$, $p < .05$. Even though there is a confounding of readability and style, these results tentatively support the conceptualization of this difference between expository and narrative styles. This conceptualization needs further testing utilizing materials of equivalent readability. Additionally, it is felt that this type of analysis may prove useful in future cloze research.

Table 2

Mean Percent of Correct Responses
on One-Fifth Cloze Passages
in Three Item Categories

	CTX 1	CTX 2	CTX 3	TOTAL *M*
Kon-Tiki (Nar)				
Total	39.48	42.26	41.42	41.05
Lexical	17.26	20.00	20.48	19.25
Function	54.30	57.10	55.38	55.59
Body Language (Exp)				
Total	36.00	36.77	33.38	35.48
Lexical	22.70	23.30	18.88	21.63
Function	51.61	52.60	51.05	51.75

Traditionally, it has been demonstrated that cloze correlates highly with standardized reading comprehension tests. When one wishes to examine specific aspects of the comprehension process, cloze seems less appropriate as a measure, since it seems to force processing to occur at the sentence level. The fact that the deletion patterns have been based on statistical, not linguistic models may be another reason why cloze has been insensitive to the natural cuing systems in connected discourse.

Bowers and Nacke (1971-72) and Ohnmacht, Weaver and Kohler (1970) have proposed that deletion patterns be based on a linguistic model. One promising approach to determine those linguistic items which provide cohesion in discourse is that of Finn (1977-78). Once such items have been identified, deletion patterns based on these items can be constructed as more accurate measures of contextual buildup. Additional factors which must be considered in developing new deletion patterns are those associated with the differences in style and organization of textual material.

References

Aborn, M., Rubenstein, H. & Sterling, T. Sources of contextual constraint on words in sentences. *Journal of Experimental Psychology*, 1959, *57*, 171-180.

Bowers, F. & Nacke, P. L. Cloze, transformational theory and redundancy. *Journal of Reading Behavior*, 1971-72, *4*, 20-33.

Bransford, J. D. & Johnson, M. K. Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 717-726.

Fillenbaum, S., Jones, L. V. & Rapoport, A. The predictability of words, and their grammatical classes as a function of rate of deletion from a speech transcript. *Journal of Verbal Learning and Verbal Behavior*, 1963, *2*, 186-194.

Finn, P. J. Word frequency, information theory, and cloze performance: A transfer feature theory of processing in reading. *Reading Research Quarterly*, 1977-78, *13*, 508-537.

Goodman, K. S. Reading a psycholinguistic guessing game. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (2nd ed.). Newark, Delaware: International Reading Association, 1976.

Iowa silent reading tests. New York: Harcourt Brace Jovanovich, Inc., 1972.

MacGinitie, W. H. Contextual constraint in English prose paragraphs. *Journal of Psychology*, 1961, *51*, 121-130.

Menosky, D. M. *A psycholinguistic description of oral reading miscues generated during the reading of varying proportions of text by selected readers from grades two, four, six and eight.* Unpublished doctoral dissertation, Wayne State University, 1971.

Miller, G. R. & Coleman, E. B. A set of thirty-six prose passages calibrated for complexity. *Journal of Verbal Learning and Verbal Behavior*, 1967, *6*, 851-854.

Ohnmacht, F. W., Weaver, W. W. & Kohler, E. T. Cloze and closure: A factorial study. *Journal of Psychology*, 1970, *74*, 205-217.

Ramanauskas, S. O. The responsiveness of cloze readability measures to linguistic variables operating over segments of text longer than a sentence. *Reading Research Quarterly*, 1972, *8*, 72-91.

Rummelhart, D. E. *Toward an interactive model of reading* (Tech. Rep. No. 52). La Jolla, California: University of California, San Diego, Center for Human Information Processing, 1976.

Sager, N. Informal presentation. Psycholinguistic Circle, New York University, Spring 1977.

Schallert, D. L. *Improving memory for prose: The relationship between depth of processing and context* (Tech. Rep. No. 5). Urbana, Illinois: University of Illinois, Laboratory for Cognitive Studies in Education, 1975.

Smith, F. *Comprehension and learning*. New York: Holt Rinehart & Winston, 1975.

Taylor, W. L. Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, 1953, *30*, 414-433.

Taylor, W. L. Recent developments in the use of cloze procedure. *Journalism Quarterly*, 1956, *33*, 42-48.

GLENN M. KLEIMAN
DIANE L. SCHALLERT
University of Illinois
at Urbana-Champaign

Some things the reader needs to know that the listener doesn't

Much attention has been paid to the similarities between listening and reading. In both teaching programs and research, it has often been assumed that reading comprehension ability consists of a simple combination of listening comprehension and print-to-sound decoding abilities. Clearly there are many important similarities between listening and reading, and through experiences with spoken language the child acquires much of what he needs to know in order to read. However, this emphasis on the similarities between listening and reading has led to the neglect of some important differences. This paper will focus on these differences and consider the question: What do skilled readers need to know that they would not have acquired via experience with spoken language? Particular emphasis will be placed on the listening and reading tasks encountered by transitional readers, those who have mastered individual word decoding but still have a lot to learn before they can fluently comprehend written language.

The differences between the written and spoken language school children encounter will be divided into five categories: (1) differences due to the existence of intonation and stress in speech but not in writing; (2) differences in the situations in which speech and writing are generally used; (3) differences in the functions speech and writing most commonly serve; (4) differences in characteristics such as vocabulary, syntactic complexity, and amount of redundancy; and (5) differences due to the permanent nature of writing and the transient nature of speech. In the final section, we will consider the ways in which these differences interact to make the task faced by novice readers different from the listening tasks with which they are already familiar.

Intonation and Stress

It is obvious that speech contains intonation and stress while writing does not. However, it may not be immediately obvious that intonation and stress provide information that is useful to listeners. Intonation helps listeners divide the stream of speech into meaningful word groups. Stress is useful in determining which is the new or important information in a sentence.

First, consider the use of intonation to help determine which words go together to form a meaningful group. Such groups of words are sometimes called constituents. There is a large variety of evidence that such grouping is necessary to understand any but the simplest sentences, whether they are presented in writing or in speech (e.g., Clark & Clark, 1977). Consider the following sentence (Graf & Torrey, 1966) divided in two different ways:

A	*B*
During World War II even fantastic schemes received consideration if they gave promise of shortening the conflict.	During World War II even fantastic schemes received consideration if they gave promise of shortening the conflict.

A is easier to read than *B* because it is divided at constituent boundaries. That is, it has been prechunked into meaningful units. Now consider reading this sentence aloud. Clearly the pauses tend to fall along the breaks given in *A*, not those given in *B*. This is a well documented finding: Pauses and other intonation features often provide cues to constituent boundaries.

The reader also must determine which words go together to form meaningful groups. How can this be done without intonation? Punctuation provides some help, but is generally not sufficient; most constituent boundaries are not marked by any punctuation. The reader must depend on syntactic and semantic cues to constituent boundaries. For example, some types of words, such as determiners (a, the), quantifiers (some, all, many), and definite pronouns (I, you, she) usually occur at the beginning of a constituent. The use of syntactic cues such as these seems to require more complex knowledge and processing than the use of intonation cues; they simply are not as obvious as intonation. This may cause difficulty for novice readers in determining the constituents of written sentences, and therefore in comprehending them.

Stress provides cues useful in separating the new or focal information contained in a spoken sentence from the less important information (Bolinger, 1972). For example, consider the following sentences spoken with the capitalized word stressed:

C. JOHN stole the picture.
D. John STOLE the picture.
E. John stole the PICTURE.

In each case the stressed word would be the one carrying the new information. That is, sentences C, D and E could be answers to questions C′, D′ and E′, respectively.

C′ Who stole the picture?
D′ What did John do with the picture?
E′ What did John steal?

The stress on the new or important terms provides the listeners with useful information. Readers must compensate for the lack of stress cues and are forced to rely upon less obvious cues to identify the important information. Readers must make greater use of syntactic cues (compare *It was John who stole the picture* with *It was the picture that John stole*). Also, readers must make greater use of information from previous parts of the text and from their own knowledge to determine which is the important or new information. Again, the lack of cues which are available in speech forces readers to use more complex knowledge and processes than listeners, and this may present problems for novice readers.

Differences in the Situations in which Speech and Writing are Used

Anything written can be read aloud and anything spoken can be written down. However, the two modes are by no means interchangeble. Some situations and purposes call for spoken communication and others for written. For example, speech is most commonly found in situations where the communicants are in the same place. Therefore, speakers and listeners often share a mutual nonlinguistic context and are able to interact with each other. Writers and readers generally do not share a common nonlinguistic context and generally cannot interact with each other. A shared context would often facilitate understanding and we know that young children rely a great deal on context to help them understand speech. Also, when the communicants can interact the speaker can take into account what the listener already knows and doesn't know, and the listener can direct the speaker by asking questions, making comments, or just looking puzzled. These differences are considered further in the paper by Rubin in this volume.

Differences in the Functions of Written and Spoken Communication

Speech and writing also differ in their usual functions. There is a strong tendency for speech to be used for informal social communications and writing for the more formal communication of information. This difference may result in difficulties for novice readers in two ways. First, the reading tasks they face may often assume knowledge that would not be necessary to understand the spoken language they usually encounter. That is, the acquisition of much new knowledge and the extension of existing knowledge is necessary for successful reading. Secondly, interpersonal communication may be much more motivational than informational communication. Many children may lack motivation to work at understanding the abstract, formal, detailed language often found in writing.

Differences in the Language Used in Speech and Writing

Studies have found that the actual language used in writing tends to differ from language used in speech in a variety of characteristics. DeVito (1965) compared samples of the writing and speaking of ten speech professors on topics of professional interest. He found the writing contained longer and less common words, as well as a larger diversity of words. Similar studies have found that writing tends to be less redundant than speech: Speakers often repeat themselves, either verbatim or in paraphrase (Walker, 1975; Horowitz & Newman, 1964; Wilkinson, 1971). Furthermore, it has been proposed that writing tends to be syntactically more complex (as indicated, for example, by frequency of subordinated and conjoined clauses) and more detailed and precise than speech (Horowitz & Berkowitz, 1967; Wilkinson, 1971). It has also been suggested that certain types of complex discourse structures may be more natural in writing (Danks, 1974).

If these differences hold for the speech and writing children encounter, they would entail differences in the knowledge necessary for successful reading and listening. The novice reader may well face more complex vocabulary, sentence syntax, and discourse structures than he had previously encountered in speech, and therefore would need to extend his knowledge in these areas. He also must adjust to the greater detail and precision found in writing, and to learn to take advantage of the permanence of writing to compensate for its lack of repetition.

Using the Permanence of Written Text

The permanence of writing provides readers with some very useful options not available to listeners. Readers can proceed at their own rate while listeners must follow the material as the speaker presents it (although this may often be compensated for by the option of interacting with the speaker). Furthermore, the pace is not simply set and then maintained throughout the text. Skilled readers slow down for important or confusing passages and speed up for easy or unimportant ones. Another option available to readers is returning to previously read parts of the text. Skilled readers do this often, going back to reread as little as a single word or phrase or as much as a large section of text. Efficient readers may also take advantage of the permanence of writing by previewing the text to organize further reading (Robinson, 1970). In general, skilled readers take advantage of the permanence of written text by efficiently extracting the information they need as they need it, but poor readers may not do so (Neville & Pugh, 1976-1977).

Efficient extraction of information from text requires some skills which have been generally neglected in studies of reading. Readers need to monitor their own comprehension so they can determine when rereading is necessary and they need to evaluate what they are reading to determine if it is important and needs to be read slowly and carefully. Very little is known about how skilled readers do this monitoring and evaluation or about how these skills develop, but the limited available evidence suggests that monitoring and evaluating may be surprisingly difficult for young children (Markman, 1977; Brown & Smiley, 1977). The importance of these skills in reading, their development, and how they can be trained are clearly in need of further study.

Interactions of the Differences Between Written and Spoken Language

Differences from these five categories interact in determining how the skills and knowledge necessary for successful reading differ from those necessary for successful listening. By way of summary, some of the areas in which the novice reader may need to acquire new skills and knowledge will be reviewed.

Successful reading may require more comprehensive general knowledge of the world than listening for a number of reasons. Writers generally cannot tailor their message to fit a particular reader, while speakers often can. Also, writers are unable to receive continuous feedback from the recipients of their message, and are not available to answer requests for clarification, as speakers are in many situations. As Socrates tells us in the Platonic dialogue *Phaedrus*: "Written words seem to talk to you as though they were intelligent, but if you ask them anything about what they say . . . they go on telling you the same thing forever."

Since readers are unable to influence how the message is communicated, they must depend upon their own abilities and knowledge to interpret it. The use of writing for more informational, rather than interpersonal, communication, and the greater detail and precision found in writing, also contributes to the novice readers' need to increase and expand their knowledge.

Novice readers also need to increase their knowledge of syntax and vocabulary over that acquired via listening. The syntax encountered in writing may often be more complex, and the vocabulary more diverse, than that found in speech. Also, since intonation and stress are not available, readers must depend more upon syntactic and semantic cues to determine constituent boundaries and which terms convey the focal information. Furthermore, readers must comprehend the syntax and vocabulary as it is written; they cannot interact with the writer to ask for clarification and they are less likely than listeners to have the same information restated in a different form.

Although writing presents some unique difficulties, it also provides the reader with some options that, when used properly, can facilitate comprehension and perhaps compensate for some of the ways it is more difficult than listening. Since writing is permanent, readers can set their own pace, reread when necessary, and preview the material to organize further reading. Efficient use of these sampling options requires that readers monitor their own comprehension, so they know when to reread or slow down, and evaluate what they are reading, so they can attend carefully to the material that is important for their purposes. These two skills of monitoring and evaluating may be very difficult for novice readers.

We have described several types of knowledge that novice readers may need to acquire or increase, and several types of cognitive processing they need to master. It is important to realize that readers cannot simply deal with one of these requirements at a time, but must use all these types of knowledge and processes at once. Even if a reader is capable of monitoring his or her own comprehension, evaluating the material, using syntactic and semantic cues to determine constituent boundaries and the important terms, understanding the vocabulary and syntax, and using the required general knowledge, doing all of these at once may overcome his or her attention and processing capacities. That is, even with each individual component mastered, combining them into efficient reading may present difficulties.

References

Bolinger, D. Accent is predictable (If you're a mind reader). *Language*, 1972, *48*, 633-644.

Brown, A. L. & Smiley, S. S. Rating the importance of structural units of prose passages: A problem of metacognitive development. *Child Development*, 1977, *48*, 1-8.

Clark, H. H. & Clark, E. V. *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt, Brace & Javonovich, 1977.

Danks, J. H. Comprehension in listening and reading: Same or different? *Proceedings of the Interdisciplinary Institute in Reading and Child Development*. University of Delaware, 1974.

DeVito, J. A. Comprehension factors in oral and written discourse of skilled communicators. *Speech Monographs*, 1965, *32*, 124-128.

Graf, R. & Torrey, J. W. Perception of phrase structure in written language. *APA Convention Proceedings*, 1966, 83-84.

Horowitz, M. W. & Berkowitz, A. Listening and reading, speaking and writing: An experimental investigation of differential acquisition and reproduction of memory. *Perceptual and Motor Skills*, 1967, *24*, 207-215.

Horowitz, M. W. & Newman, J. B. Spoken and written expression: An experimental analysis. *Journal of Abnormal and Social Psychology*, 1964, *68*, 640-647.

Markman, E. M. Realizing that you don't understand: A preliminary investigation. *Child Development*, 1977, *48*, 986-992.

Neville, M. H. & Pugh, A. K. Context in reading and listening: Variations in approach to cloze tasks. *Reading Research Quarterly*, 1976-1977, *12*, 13-31.

Robinson, F. A. *Effective study*. New York: Harper & Row, 1970.

Walker L. Comprehending writing and spontaneous speech. *Reading Research Quarterly*, 1975-1976, *11*, 144-167.

Wilkinson, A. *The foundations of language: Talking and reading in young children*. London: Oxford University Press, 1971.

ANDEE RUBIN
Bolt Beranek and Newman Inc.

A taxonomy of language experiences

Children come to the task of reading with a set of well-practiced oral language comprehension skills. This linguistic skill, which a child has developed in a remarkably short time, obviously greatly facilitates learning to read. In fact, Huey (1908) regards reading as a process of decoding orthographic symbols to a phonemic representation, then comprehending that as if it were speech: "The child comes to his first reader with his habits of spoken language fairly well formed and these habits grow more deeply set with every year. His meanings inhere in this spoken language and belong but secondarily to the printed symbols." Yet the current furor over children's inability to comprehend what they read indicates that a child's oral language skills in and of themselves must not be sufficient precursors to reading. Although it is clear that the necessity for visual decoding is a difference between children's oral language and reading comprehension, it is but one of a great many distinctions, all of which may well present stumbling-blocks for children learning to read. This paper introduces a taxonomy of the differences between children's typical oral language experiences and the experience of reading a book. By delineating the dimensions along which these language experiences differ, we hope to understand better the nature of the cognitive leap we expect children to make in learning to read. In addition, the set of differences between oral conversations and written text gives us the tools to specify the relationships among other language experiences such as watching television or plays, talking on the telephone and reading comics and to understand the cognitive demands each of these experiences makes on the child as comprehender.

The majority of a child's oral language experiences may be described as interactive conversations in which the child participates as both speaker and listener. All the participants share a spatial, temporal and situational context and their verbal communication is augmented by intonation, facial expression and gestures. The differences between this situation and that of a child reading a story may be divided into two large sub-categories: those having to do with the communicative *medium* and those dealing with the *message*; each of these subcategories is further divided into dimensions. The emphasis here will be on *medium* dimensions and on the new demands these differences impose on children learning to read; the message dimensions — topic, structure and function — will be discussed only briefly, but a more complete description may be found in Rubin (1978).

There are seven dimensions along which the communicative medium of a language experience can be placed. If we think of a space defined in terms of these seven dimensions, a child's oral language experience, as described above, would lie on the opposite end of a long diagonal from reading stories, with one point being (0,0,0,0,0,0,0), the other (1,1,1,1,1,1,1). The medium related dimensions are: modality, interaction, involvement, spatial

4

commonality, temporal commonality, concreteness of referents, and separability of characters. While these distinctions are in some measure intuitive, there is as yet little or no evidence that any one of them is a crucial difference which causes problems for children learning to read. The requisite experiments have not been done at least in part because these dimensions have not been succinctly identified before. The framework presented here should be regarded, then, as a set of hypotheses and suggestions for future research. Further descriptions of the dimensions are as follows:

1. MODALITY — is the message written or spoken? This dimension is the one on which most research on the relationship between listening and reading has focused (see Danks, 1977, for a review of this work). In fact, that work has mainly concentrated on only one aspect of this distinction: the added necessity of visual decoding in reading. Even in this single dimension, however, there are other differences which impinge substantially on the processing demands of the comprehension task.

Spoken language has as one of its most salient aspects the use of stress, intonation, and other prosodic features. Temporal characteristics of speech such as pauses and changes in speed often provide clues for the chunking of words into larger constituents, while stress indicates discourse organizing concepts such as the distinction between given and new information (Clark & Clark, 1977). The transition to text requires the development of alternate strategies to compensate for the disappearance of these features.

Text does have some compensatory aspects, however. Punctuation provides some of the clues for which prosodic features are useful. In contrast with speech, segmentation of the message into words and sentences is concretely indicated in written text and is not a task which must be performed by the reader. In addition, certain devices which are used solely in text such as paragraphs, underlining and italicizing can help specify the larger structure of the message. Effective reading involves recognizing the function of these aids and developing processes to take the best advantage of them.

Another characteristic of text which can be an asset in its comprehension is its permanence. Readers can use this fact by looking back over passages they have previously read, re-reading a sentence which was misparsed the first time around or re-reading an entire paragraph whose point became clear only at the last sentence. A major strategy a child must develop in making the transition from oral to written language is a method for using the permanence of text to compensate for some of its differences from speech.

2. INTERACTION — is the hearer/listener able to interact with the speaker/writer? Clearly, in a conversation, each participant has a chance to speak and often uses this opportunity to indicate that he or she has not understood the speaker. Thus, in a conversation which is "working," the hearer can verify his or her hypotheses quickly, making the maintenance of competing hypotheses less necessary. The parallel construction of hypotheses is a skill which we hypothesize children must learn in their transition to reading stories.

Being in a conversation also requires the listener to make an active attempt to understand what is being said in order to respond appropriately. In non-interactive media such as books and TV this impetus is absent. Participatory language experiences are, in addition, highly individualized; each participant has some model of the other's beliefs and knowledge and composes utterances taking this model into account. Thus, the language with which a child comes into contact in conversatons is more tailored to his or her knowledge than the language in a multi-recipient object like a book could be. Even in the best of situations, then, a child reading a book will encounter unfamiliar material more often than in having a conversation with a parent or peer.

3. INVOLVEMENT — is the communication directed to the reader/listener? The inclusion of this dimension reflects the fact that certain language experiences are directed toward the reader/listener, while in others, he or she is essentially "eavesdropping." One clue to locating a language experience along this dimension is the use of second person pronouns. An "involving" communication will use "you" to refer to the reader/listener, sometimes even in the imperative. If a "noninvolving" communication contains "you" at all, the referent will be a character in the story, or a generalized person ("You never know what's going to happen next."). Involvement in a communication act usually implies that the writer/speaker knows who the reader/listener is; consistent with this implication is the fact that most written communications of this sort are derived from oral situations (e.g., letters).

4. SPATIAL COMMONALITY — do the speaker and listener (reader and writer) share a spatial context? This dimension really comprises two different questions. The first might be phrased: Can the participants see one another? The second: Can the participants use the same spatial deictic terms bcause they are in the same place?

The first question is primarily one of extra-linguistic communication. Gestures, facial expressions and pointing can all be used to facilitate communication. A nod of the head may denote agreement; a puzzled look may communicate a lack of understanding, causing the speaker to restructure the utterance. Pointing may aid in specifying referents for pronouns or noun phrases such as "that dog over there."

The second aspect of spatial commonality has to do with the use of deictic words such as "here," "there," "come," "go," etc. If the two participants are in the same place, they can understand such words without translating them to account for the other person's being in a different place. The permanence of written language and the existence of modern telecommunications have created situations in which the two participants can be separated in space, thus making it necessary for the listener to interpret spatial deictic terms in the speaker's context. In related work both Tanz (1976) and E. Clark (1977) have noted that children's ability to understand spatial deictic terms even in oral language is not complete when they enter school.

5. TEMPORAL COMMONALITY — do the participants share a temporal context? This again is a deictic issue involving the use of such words as "now," "today," "last Sunday," and verb tense markers. The correct interpretation of such words when the participants are separated in time requires the reader/listener to take the point of view of the speaker/writer. A child's oral language experience does not often require this ability to switch the temporal context of utterances. Although it is certainly possible for a mother to address the following remark to her child: "Remember I told you yesterday, 'You can go out to play tomorrow'.", it appears that this demand for temporal context-switching is seldom imposed on a child in oral conversations.

6. CONCRETENESS OF REFERENTS — are the objects and events referred to visually present? Early conversations deal almost exclusively with concrete objects which a child can see: Mommy, Daddy, clothes, food, or objects which the child has at least seen previously and which therefore have some concrete reality to him or her (Nelson, 1974). In reading or listening to stories, a child is

often required to make up an object or event given an incomplete, verbal description, a process which may take additional cognitive sophistication. The child may also have to integrate several partial descriptions of the same object and remember the composite description without the aid of an external referent.

7. SEPARABILITY OF CHARACTERS — is the distinction between different people's statements and points of view clearly indicated? In a normal conversation, such distinctions are obvious, as each person makes his or her own statements; each point of view has a physical "anchor." Even so, for a young child, the parallel maintenance of several distinct points of view may be confusing. In a book this problem is compounded, as the child must not only "construct" the individuals involved (see concreteness of referents, above), but must parcel out comments, feelings and motivations to each of them on the basis of more subtle clues: punctuation, paragraph structure and inferences based on some consistent model of each of the characters.

These seven medium dimensions, however, account for only some of the differences among language experiences. Conversations and texts diverge significantly in terms of the structure, function and topic of the message itself. The *structure* of most conversations (short statements related as question/answer or on a common topic) is a far cry from that of stories (episodic structure, including setting, character introduction and action (Rumelhart, 1975)). Differences in vocabulary and syntactic structure between oral and written language have also been noted (Danks, 1977). In terms of *topic*, many of a child's conversations center around concrete, familiar objects; the transition to reading involves a move toward more abstract subjects often unfamiliar to the reader. The *function* of reading also differs dramatically from that of having a conversation. Children usually engage in conversations to get information, cause an event or make social contact. The function of texts tends more toward informing, persuading, or, in the case of early-grade readers, teaching new words. The purpose of these texts often escapes children. They demand, "Why should I read?"; interestingly enough, they never ask, "Why should I talk?".

Although these dimensions have been identified and discussed by contrasting two extremes — children's oral conversations and reading a story — there are many language experiences which lie between the two. A dimensionalization like the one presented here defines a space within which language experiences may be compared and inspires a search for the uninstantiated possibilities. We can think of each language experience to be described as a point in 7-dimensional space. At first the space appears to be only sparsely filled, but, in fact, we can come up with quite a few intermediate points by teasing apart the dimensions listed above.

Take, for example, talking on the telephone, a language experience which is very similar to a face-to-face conversation except for the lack of spatial commonality between participants. This difference implies that a child talking on the telephone faces the potential problem of incorrectly interpreting words such as "here" because of the spatial context shift necessary to interpret the word. An additional hindrance implied by the position of this experience on the "spatial commonality" dimension is the lack of extralinguistic communication, made impossible by the limited communicative medium. Objects referred to in the conversation which are in the speaker's spatial context are probably not immediately visible or accessible to the child. For a child who relies on these aids to comprehend speech, their absence may necessitate additional processing and/or lead to comprehension difficulties. The point is that this additional processing is precisely the type which is necessary in reading stories as well.

Another example is reading comics. Although this activity shares with reading stories the act of decoding, it differs from it along several dimensions. Characters in comic books are easy to identify because they are visually represented and their words and actions are clearly attributed to them. The multitude of pictures also makes it easier for a child to visualize referents (animals, castles or dragons) and make it less necessary for him or her to integrate successive partial descriptions into a larger one, as the picture provides many of the details at first glance. The structure of the text in comic books tends to be conversational, so they may lie closer to certain oral language experiences on the message dimensions (structure, function and topic) as well.

The existence of this space of language experiences, then, allows us to make hypotheses about some distinctions between oral conversations and written text which may really make a difference to children learning to read. It also provides a framework for identifying how other language experiences might be related to conversing and reading, what cognitive demands they share with each and how we might use them to ease the transition for children.

References

Clark, E. V. From gesture to word: On the natural history of deixis in language acquisition. In J. S. Bruner & A. Garton (Eds.), *Human growth and development: Wolfson College lectures 1976*. Oxford: Oxford University Press, 1977.

Clark, H. H. & Clark E. V. *Psychology and Language*. New York: Harcourt Brace Jovanovich Inc., 1977.

Danks, J. H. *Comprehension in listening & reading: Same or different?* Submitted for the report of the Interdisciplinary Institute in Reading & Child Development, University of Delaware, June-July 1977.

Huey, E. B. *The psychology and pedagogy of reading*. Cambridge, Mass.: MIT Press, 1968.

Nelson, K. Concept, word and sentence: Interrelations in acquisition and development. *Psychological Review*, 1977, *81*, 267-285.

Rubin, A. D. A theoretical taxonomy of the differences between oral and written language. To appear in R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical Issues in Reading Comprehension*. Hillsdale, N.J.: Lawrence Erlbaum Associates. Also as Center for the Study of Reading Technical Report No. 35, 1978.

Rumelhart, D. E. Notes on a schema for stories. In D. Bobrow & A. Collins (Eds.), *Representation & understanding: Studies in cognitive science*. New York: Academic Press, 1975.

Tanz, C. *Studies in the acquisition of deictic terms*. Unpublished doctoral dissertation, University of Chicago, 1976.

EARL F. RANKIN
University of Kentucky

Characteristics of the cloze procedure as a research tool in the study of language

The purpose of this paper is to demonstrate the value of the cloze procedure in developing comprehension measurements for both written and oral language research. It will be argued that the cloze procedure is uniquely qualified for this important function due to (a) its theoretical base in an information processing theory of learning and communication and (b) its superior psychometric potential for generating valid tests of language comprehension.

For the past twenty years, the writer has been collecting bibliographical references and documents in which the cloze procedure has been used as a research tool for studying a wide variety of language variables. At the present time, approximately six hundred references have been gathered. Although the major thrust of numerous papers on the cloze procedure delivered at annual meetings of the National Reading Conference has been on the measurement of reading comprehension and the readability of printed passages, there has been a steady growth in the literature on many other facets of language, both written and spoken. A rough count of titles in this collection indicates that studies of language and cognitive variables constitute the largest single category in the bibliography. In fact, the number of language studies using the cloze procedure is twice as great as those references in which the measurement of reading comprehension or readability is the major focus. Many language topics have been studied, such as grammar, dialect, oral speech characteristics, foreign language facility, listening, contextual constraints in prose, communication effectiveness, drug effects on communication and cognitive processes, divergent and convergent thinking, etc. As the psychology of reading has moved into the mainstream of language and/or cognitive theory and research and out of the shallow waters of educational methodology, the cloze procedure has served as a powerful research tool in facilitating this rapprochement. This development has no doubt been made possible by the fact that the cloze procedure was derived from an information processing theory which is equally applicable to all forms and modes of communication. The increasing use of this technique in language investigations also stems from its superior psychometric characteristics.

Theoretical Foundations of Cloze Measurements

It is axiomatic among measurement specialists that most measuring instruments in psychology and education are not firmly rooted in theory. Reading tests, for example, are usually based upon an empirical set of "skills" stemming from traditional classroom usage. Skills to be measured are selected more or less pragmatically in an effort to provide useful information for teachers and clinicians. Like most tests, measures of language comprehension are largely based upon the shifting sands of empiricism. It is the writer's opinion that the cloze procedure produces superior language tests because it has, potentially, better construct validity than conventional measurement procedures.

As we all know, the term "cloze" was coined by Wilson Taylor from the term "closure" in gestalt psychological theory. However, Taylor did not attempt initially to work out a careful and detailed theoretical analysis of the cloze procedure in relation to the gestalt theory of perception and learning. Despite the use of the term "cloze," Taylor's theoretical derivation of the cloze procedure was essentially based upon a modification of a mathematical "information theory" designed by Shannon and Weaver (1949). Two basic concepts in this theory were "information" and "redundancy." Information was defined in terms of how much uncertainty was reduced by the accurate reception of a message and redundancy was defined in terms of information coming from more than one source. It is important to note that the cloze procedure was initially derived from a practical engineering strategy for measuring the flow of information. This accounts both for its pragmatic appeal and its theoretical significance to increasing numbers of scholars interested in language and cognitive processes.

It remained for Smith (1971, 1975) to work out more fully a psycholinguistic theory of comprehension which made extensive use of the concept of information processing. It is the information processing aspect of Smith's rather complex conceptualization of comprehension which is particularly relevant to a fuller understanding of the construct validity of the cloze procedure as a measure of information processing. Smith viewed individual differences in ability to comprehend among readers or listeners in terms of their ability to reduce uncertainty through the use of both distributional and sequential redundancy. The message receiver continually makes inferences concerning the information to be supplied by filling in gaps at choice points throughout a message. Thus, the process of inference is a basic component of all comprehension. Filling in inferential gaps reduces uncertainty and thereby provides continuing feedback as the comprehender actively participates in the communication process. This theoretical portrayal of comprehension by a reader or listener as a processor of information provides a solid psycholinguistic basis for using the cloze procedure in constructing tests of language comprehension.

Let us consider, hypothetically, the process by which a cloze response might be made in terms of the previously described theory of information processing, assuming a high level of uncertainty at a cloze gap. In taking a cloze test, as in all normal reading and listening, the individual reacts to information from several sources (i.e., redundancy) in making inferences among words to be considered for filling in gaps within a message. In the written or spoken sentence, "The cat climbed ____________ the tree in order to escape from the dog", the comprehender would react to the redundant information stored within his nervous system together with visual and/or auditory information in the message in making a cloze response. Previous experience, either direct or vicarious, about dogs chasing cats and cats climbing trees would provide a greater source of information than that found on the page or in the spoken message. The comprehender might react first to the sentence structure and eliminate such possible words as nouns, verbs, adjectives, and adverbs as alternatives to be considered within this context. By making use of sequential redundancy concerning grammatical patterns, he would reduce considerable uncertainty because he has only the category of function words to be considered as possible choices. Next, the comprehender might make use of distributional as well as sequential redundancy in making inferences. Among function word alternatives, certain words are more likely to occur than others independently of verbal sequence. For example, the function words "up" or "down" are statistically more likely to be used than the function word "throughout", not considering their context. The skilled comprehender has both sequential and distributional redundancy "rules" programmed in his nervous system.

Semantic clues would also suggest that the word "up" would be a better choice than the word "down" as a more likely means of escape from the dog. Other overlapping sources of information used in reducing uncertainty at the cloze gap might be visual-orthographic and/or auditory-phonemic sequences. The skillful comprehender unconsciously makes efficient use not only of multiple sources of information but also of overlapping types of redundancy to reduce uncertainty at cloze gaps randomly distributed throughout a message. In some messages, the sophisticated message receiver could use elements of the author's writing style and even aesthetic elements in making cloze inferences.

It is important to realize that all readers and listeners carry out some such process of inference and redundancy utilization as previously described in comprehending messages. We only tend to become aware of this inferential process when there is a high level of uncertainty at a gap. The main difference between taking a cloze test and reacting to a normal written or spoken message is that in the real world the reader or listener selects his own gaps in relation to his own background of information, his current motivational state, temporary fluctuations in his attention, etc. In the cloze message, the gaps are arbitrarily chosen by the test constructor. However, any test situation is necessarily somewhat artificial. The writer believes that this model of information processing is valid and supported by a growing body of research and theory. The "goodness of fit" between this model and the cloze procedure is impressive and very significant to the serious language researcher.

Unique Measurement Characteristics for Language Studies

Let us consider characteristics of cloze measurements from the standpoint of measurement theory. At first blush, a cloze measuring instrument for assessing comprehension is not very convincing. The naive observer might ask, "How could anyone reasonably be expected to fill in all of those blanks?" After all, a cloze test looks (or sounds) much like the type of fill-in-the-blank exercises used to measure rote learning in many workbooks. Moreover, sophisticated critics have maintained that, on its face, a cloze test cannot conceivably measure comprehension directly. It is charged that if cloze tests correlate highly with other comprehension measurements, then surely this must be due to the fact that both the cloze and the comprehension tests are measuring some common variable.

Now there is no doubt that the cloze procedures tests which are lacking in face validity. A cloze test does not look or sound like a normal communication, but the previous analysis of comprehension in terms of information processing theory should be sufficient to dispense with the reactions of the naive observer to a cloze measurement. However, the criticism that the cloze test does not measure comprehension directly but measures some other variable which is correlated with comprehension must be taken more seriously.

It has been asserted that the correlation between cloze measurements and other comprehension measurements is due to the fact that both are measuring a general verbal competency. There is no doubt some substance to this contention. High correlations are usually found between cloze test scores and measures of verbal aptitude. However, this is also true of other comprehension tests. In any event, the correlations between both cloze test results and conventional comprehension results on the one hand, and verbal aptitude tests on the other, do not account for a large amount of the variance in common to both distributions. Therefore, cloze tests measure something more than verbal aptitude.

A more serious criticism has been made by Carroll (1972) to the effect that cloze scores are largely influenced by linguistic clues in the immediate context around the missing word and therefore, cloze scores do not assess the ability to comprehend major ideas in a message. If this were the case, the use of cloze tests as comprehension measures, per se, would be questionable indeed. This criticism, together with a finding by MacGinitie (1961) to the effect that context of more than five words around a blank in a printed cloze test does not help in making the correct response, has been widely quoted. However, it should be noted that MacGinitie also pointed out that constraints may operate over longer distances, since knowledge of the topic might have an influence which could extend beyond local contextual boundaries. The latter statement is important to keep in mind. Since the cloze procedure measures language correspondence between a message source and receiver, redundancy resides in both syntactic and semantic information within the shared language correspondence. This is obviously not restricted by localized visual or auditory contextual clues.

Studies by Darnell (1963), Ramanauskas (1971), and others tend to show that cloze responses are indeed sensitive to longer range constraints than reported by MacGinitie. These studies show that scrambling sentence order produces a significant influence on cloze responses within sentences. This could hardly be the case if cloze tests measure only "local redundancy." This issue is not by any means "clozed," but the writer is convinced that the evidence does not at this time give strong support to Carroll's position.

From a theoretical point of view, a good case can be made for the contention that cloze measurements do in fact measure comprehension more "directly" than conventional comprehension tests. Five points will be made to support this contention. First, cloze tests are intrinsic measures of the effectiveness of communication by sampling the degree of language correspondence between a message source and a receiver. Substantially the same results are obtained whether the scoring is done by exact word method or by the synonym method. This could hardly be the case if comprehension of the communication were not being tapped directly. Second, cloze tests measure comprehension *in process*, not comprehension as a product after the fact. Answering large numbers of questions after the communication has been received (as in the conventional comprehension test) is not as direct a measure of the communication in process as can be obtained by the cloze test. Third, all cloze responsed are based upon the basic psycholinguistic process of inference which is intrinsic to all communication. We avoid overloading our short-term memory by tuning in and out selectively and filling in the gaps in both oral and written communications. It is precisely this process that is tapped by all cloze items in varying degrees. This cannot be said of conventional comprehension measurements. Fourth, cloze tests sample more or less randomly the choice points for predictability within a message. What other comprehension test can attain such unbiased item samples from a universe? Fifth, unlike any other communication test format capable of measuring higher level thought processes, cloze item writing lends itself to precise replication by independent writers. Although all comprehension tests impose some degree of artificiality upon the message receiver in the measurement process, the intimate relationship between language and learning theory and cloze measurement provide a more direct and natural testing situation, in many ways, than provided by conventional comprehension tests.

Cloze measurements have been constructed and interpreted in the tradition of what has been called "classical test theory." As such, they have been designed to yield maximum variance and reliability and have received a norm-referenced interpretation. Within this theoretical framework, the cloze procedure produces excellent measuring instruments. If the goal of comprehension measurement is to obtain reliable and valid measures of even small differences in comprehension among people, cloze tests are most useful for this purpose as indicated by numerous studies of criterion related validity. The high correlations among cloze comprehension measurements for different materials written at various levels of difficulty indicate that these measurements are tapping comprehension in a very general or global sense.

Conventional cloze tests also permit a type of criterion-referenced interpretation, provided that it can be demonstrated empirically that people who fail to score above a given cut-off score cannot perform designated language tasks at a satisfactory level of competence. It should be said, however, that cloze comprehension tests do not provide criterion-referenced measurement in the sense of a test which can be interpreted in terms of mastery or non mastery of specifically defined behavioral or instructional objectives. A cloze test has proven to be a superior measure of general comprehension but does not lend itself easily to the measurement of specifically defined language comprehension processes.

In an interesting theoretical paper on reading comprehension Simons (1971) states, "The cloze test is a better measure of reading comprehension than traditional tests because it appears to measure fewer of the extraneous aspects of student functioning enumerated above. (sic) Specifically, it does not have questions and therefore is not measuring a student's skill in understanding questions. It is not a memory test because a student can continually read and examine the passage. It also does not appear to be measuring a student's familiarity with the content of the passage, at least to the degree that traditional tests do. The mechanical procedure for developing the test (the deletion of every n^{th} word) renders it more objective and less subject to the arbitrary judgments of the test constructor than traditional comprehension tests." Although Simons concludes that the cloze test of reading comprehension lacks construct validity for the same reason that all tests of reading comprehension do (the absence of a theory of the comprehension process), the writer believes that the subsequent elaboration of a psycholinguistically oriented information processing theory of comprehension by Smith provides a potentially stronger construct validity for cloze tests as measures of language comprehension than is characteristic of other comprehension tests at the present time. Empirical studies determining the construct validity of cloze tests by testing hypotheses derived from information processing theory would be desirable. Such studies would, hopefully, provide still stronger evidence for the value of cloze measurements in language research.

References

Carroll, J. B. Defining language comprehension: Some speculations. In J. B. Carroll & R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D.C.: Winston, 1972.

Darnell, D. K. The relation between sentence order and comprehension. *Speech Monographs*, 1963, *30*, 97-100.

MacGinitie, W. H. Contextual constraints in English prose paragraphs. *Journal of Psychology*, 1961, *51*, 121-130.

Ramanauskas, S. *The effect of contextual constraint beyond a sentence on cloze responses of children in special classes for the educable mentally retarded*. Unpublished doctoral dissertation, University of Connecticut, 1971.

Shannon, C. E. & Weaver, W. *The mathematical theory of communication*. Urbana, IL: The

University of Illinois Press, 1949.

Simons, H. D. Reading comprehension: The need for a new perspective. *Reading Research Quarterly*, 1971, 6, 338-363.

Smith, F. *Understanding reading*. New York, N. Y.: Holt, Rinehart and Winston, 1971.

Smith, F. *Comprehension and learning*. New York, N. Y.: Holt, Rinehart and Winston, 1975.

R. ANN ZINCK
University of the Pacific

The relation of comprehension to semantic and syntactic language cues utilized during oral and silent reading

Within the past twenty years, cloze tests have been analyzed using a variety of deletion patterns (Rankin, 1959; Weaver & Kingston, 1962; Ohnmacht, Weaver, & Kohler, 1970; Ohnmacht & Fleming, 1972; Greene, 1965; Louthan, 1965), scoring procedures, (Bormuth, 1965; Vaughn, Tierny, & Alpert, 1977), and research designs and statistical analysis (Weaver & Kingston, 1962; Bormuth, 1969; Ohnmacht & Fleming, 1970; Horton, 1975), in an attempt to describe what construct cloze measures. With the exception of the Weaver and Kingston (1962) study, research has suggested that cloze is related to a construct or constructs identified as comprehension and/or the ability to deal with the structures of language. Since research with the cloze procedure has been inconsistent in type of deletion patterns, text, age of subjects, and test criteria, a sound empirical description or theoretical framework of the process required of an individual to complete a cloze test, and consequently what is measured by cloze, has not been clearly delineated. Although variations of deletion patterns are quite common in cloze research, the scoring procedure based on the percent of correct (exact word or acceptable synonyms) responses has dominated cloze research. If it is the format which first determines the validity of cloze as an assessment instrument, then analysis of a cloze test should utilize more fully the information this format can yield. That is, cloze should be considered a medium through which one can evaluate those language processes related to the reading and completion of a cloze task. A qualitative analysis of cloze responses would seem to provide an in-depth view of the language abilities utilized in the reading of a cloze passage.

Relying upon Goodman's (1969, 1972, 1973) view of reading as a psycholinguistic process, it is then possible to see that the cloze procedure can provide "little windows" to the language-based strategies a reader uses to read a passage. Accepting the view that reading is, "one of the four language processes," (Goodman 1976a, p. 89) it thus seems sound to approach analysis of both oral and silent reading in terms of the elements of language which are integral to reading ability. According to Goodman, reading is a "spiral of predict, sample, select, guess and confirm activities" (1973, p. 295). One may thus infer that comprehending is also a continuous and simultaneous part of this activity. If reading is accepted as language process, the accusation that performance tests of reading do not reflect the reading process in relation to a model of language competence (Roush, 1976) should be examined.

At least two researchers have hinted at the feasibility and utility of coupling miscue analysis and the cloze procedure. Brown (1968) suggested that a scoring procedure which would allow a qualitative decision about the words used to complete a blank be applied to the cloze procedure. Page (1974) saw research combining the cloze procedure and miscue analysis as promising "extensive contributions to understanding the reading process" (p. 168).

Combining some of the questions raised by cloze research with the findings and theories of miscue analysis, this study attempted to examine the relation of cloze, comprehension, silent and oral reading. Miscue analysis was used to quantify syntactic and semantic language cues utilized during oral and silent reading tasks. Specifically, the questions examined in this study were:

1) How do cloze syntactic and semantic mean change scores relate to total cloze test scores?
2) How do cloze test scores and oral reading relate to a standardized test of reading comprehension?
3) How do semantic and syntactic mean change scores for content words and function words relate to literal and inferential comprehension?

Method

Subjects

The subjects for this study were 68 seventh-grade students drawn from a sample identified as reading at least one year but not more than four years below grade level, according to the *Stanford Achievement Test*. The subjects attended a middle school in a rural area of northeast Georgia.

Procedure

The Reading Comprehension subtest of the *Stanford Diagnostic Reading Test*, Level II, Form W (SDRT), (Karlson, Madden, & Gardner, 1966), was used as a criterion measure of literal and inferential comprehension.

A fifth word deletion cloze test was constructed from a story written by the researcher. The story was in a narrative style, told through the main character. No intentional controls were placed on syntax. There appeared to be a balance among simple, compound, and complex sentence structures. The plot of the story concerned a girl's attempt to establish a co-educational track team at her junior high school. The readability of the passage was estimated at upper seventh-grade according to the Readability Graph (Fry, 1968) and seventh-grade according to the Flesch Reading Ease Formula (Burmeister, 1974).

After the cloze test was constructed, the words deleted were divided into two categories; content and function words. Of the 107 total deletions, 69 were content words and 38 were function words. The KR_{21} reliability estimate of this cloze test was .73.

The cloze test was scored using three different procedures. First, the total cloze score was based on the number of exact word replacements. Second, the scoring procedures described in the Reading Miscue Inventory (RMI) (Goodman & Burke, 1972), were used in order to yield grammatical acceptability and strength of comprehension scores. The third procedure was from *The Goodman Taxonomy of Reading Miscues* (Goodman, 1976b). The "Taxonomy" was used to calculate semantic and syntactic mean change scores. These were used to analyze the degree of syntactic and semantic change for all miscues which were identified as grammatically and semantically acceptable in the RMI scoring procedure.

The third instrument was an oral reading passage which was analyzed according to procedures described in the RMI and *The Goodman Taxonomy of Reading Miscues*. Scores indicating grammatic acceptability, strength of comprehension, and syntactic and semantic mean change were obtained for the oral reading of the passage. The passage was the intact version of the cloze test. It was administered to the subjects according to the procedures described in the RMI manual.

The SDRT was administered first, followed by the cloze test. The individual administration of the oral reading passage followed one week after the administration of the cloze test.

Analysis

Pearson product-moment correlations and a forward inclusion step-wise multiple regression were used to examine the relations among the identified variables. A t-test for the significant difference between correlations was used to determine if syntactic and semantic mean change scores differentially predicted total cloze scores.

Results

Significant but small relations were found between cloze syntactic mean change scores ($r = .31, p < .01$) and semantic mean change scores ($r = .28, p < .05$) and total cloze scores. These scores did not differentially predict total cloze scores, $t(65) = .314, p > .05$.

A significant, moderate relation was found between cloze comprehension scores and the total comprehension scores of the SDRT ($r = .56, p < .01$). No stable relation ($r = -.19, p > .05$) was found between oral comprehension scores and the total comprehension scores of the SDRT.

The results of the forward inclusion stepwise multiple regression for syntactic and semantic mean change scores as predictors of literal comprehension are reported in Table 1. There was a significant, moderate relation between the two predictors, and literal comprehension. The resulting multiple R for these two predictors was .445; the R' of .415 indicated that this is a good estimate of the population correlation. Predictors three and four together added only two percent in explained variation to literal comprehension. There was a significant, moderate relation between the two predictors, cloze semantic mean change scores for content words and oral semantic mean change scores for function words, and literal comprehension. The resulting multiple R containing these two predictors was .497; and the R' of .473 indicates that this is a good estimate of the population correlation. The remaining predictors together added only three percent in explained variation to literal comprehension.

The results of the forward inclusion stepwise multiple regression for syntactic and semantic mean change scores as predictors of inferential comprehension are reported in Table 2. There was a significant but small relation ($R = .246$) between oral semantic mean change scores for function words and inferential comprehension. The remaining two predictors did not make a significant contribution to the explained variation in inferential comprehension. The fourth predictor did not meet the F-test for inclusion in the regression equation.

It appears that there is no significant relation between the syntactic predictors and inferential comprehension.

Discussion

Since cloze syntactic and semantic mean change scores did not differentially predict total cloze scores this may indicate that the subject's ability to provide responses which are syntactically similar to an expected response in a cloze blank does not provide a more reliable basis for predicting total cloze scores

Table 1

Summary of the Stepwise Regression Analysis of Predictors of Literal Comprehension

Predictor Variable	Multiple *R*	*F*
Cloze Syntactic Function Word Mean Change Scores	.394	12.097*
Cloze Syntactic Content Word Mean Change Scores	.445	8.022*
Oral Syntactic Content Word Mean Change Scores	.464	5.843*
Oral Syntactic Function Word Mean Change Scores	.466	4.350*
Cloze Semantic Content Word Mean Change Scores	.411	13.446*
Oral Semantic Function Word Mean Change Scores	.497	10.667*
Cloze Semantic Function Word Mean Change Scores	.518	7.838*
Oral Semantic Content Word Mean Change Scores	.522	5.887*

$^*p < .01$

than does the ability to provide semantically similar responses. Thus it seems that a fifth word deletion cloze test can be used to analyze an individual's ability to utilize both syntactic and semantic language cues in the completion of the cloze task. It appears that these data support Weaver's (1965) view that both syntactic and semantic constraints in the cloze procedure affect both structural and lexical language elements. The correlations in this study accounted for very little of the variation in total cloze scores, but that variation is being accounted for totally by the miscues that were either syntactically or syntactically and semantically acceptable to the entire passage. It is likely that a scoring procedure that would take into account the correct replacements would help contribute to the variance accounted for by the miscues in this analysis. Brown's (1968) view that the use of synonym scoring in cloze probably measures the reader's full language competence appears now to be somewhat simplistic. Semantic mean change scores accounted for only eight percent of the variation in total cloze. It does not seem likely that this scoring

Table 2

Summary of the Stepwise Regression Analysis of Predictors of Inferential Comprehension

Predictor Variable	Multiple *R*	*F*
Oral Semantic Function Word Mean Change Scores	.246	4.258*
Cloze Semantic Content Word Mean Change Scores	.293	3.044
Oral Semantic Content Word Mean Change Scores	.338	2.748
Oral Syntactic Content Word Mean Change Scores	.207	2.965
Cloze Syntactic Function Word Mean Change Scores	.272	2.596
Cloze Syntactic Content Word Mean Change Scores	.286	1.898
Oral Syntactic Function Word Mean Change Scores	.288	1.426

*$p < .05$

procedure alone accounts for total language competence. Certainly, degrees of semantic acceptability along with syntactic acceptability and the ability of the reader to recognize language inconsistencies and correct them enter into the evaluation of language competence.

In order to discuss the relation between cloze comprehension scores and the total scores on the SDRT, it must be noted that comprehension scores for this analysis included only those miscue patterns involving correction, semantic acceptability and meaning change which resulted in the miscue being assigned to the category of "no loss of comprehension." Since cloze miscues which indicated no loss of comprehension accounted for approximately 31 percent of the variation in SDRT scores, this finding seems to support the use of the cloze procedure combined with the miscue analysis to evaluate the reading process. The moderate relation of the cloze miscues to the comprehension subtest leads to the inference that the production of a miscue which is syntactically and semantically acceptable and with little or no loss in meaning is, at least, partially accounting for comprehension ability. Looking toward the future of cloze as a silent diagnostic test, it would be necessary to examine the role of retelling of the passage after the completion of a cloze test would play. What appeared to cloud the issue was failure to establish a stable relation between the oral strength of comprehension scores and the comprehension subtest of the SDRT. Acceptance of the concept that silent and oral reading processes are similar would lead one to expect that oral reading would also have some relation to comprehension. This discrepancy is not considered conclusive, but rather an area for further research. This finding appears to provide further empirical support for use of a retelling to evaluate comprehension following an oral assessment.

Findings concerning the ability of cloze and oral syntactic and semantic mean change scores for content and function words to predict comprehension did not reveal one test or cue system to consistently predict literal or inferential comprehension. Cloze syntactic mean change scores for function and content words were better predictors of literal comprehension than their oral reading counterparts. For semantic based language cues, cloze semantic mean change scores for content words was the first best predictor of literal comprehension, while oral semantic mean change scores for function words was the second best predictor. The single best predictor of inferential comprehension was oral semantic mean change scores for function words. The ability to perceive interrelationships between ideas (i.e., inferential comprehension) may be based on competence in dealing with structural meaning (Rankin, 1959). The ability to use this competence in oral reading appears to override the ability to use structural meaning in cloze when viewed as a predictor of inferential comprehension. However, given the mutilated characteristic of the cloze passage, it is probably more difficult for the reader to perceive overall interrelationships among the words, sentences, and ideas in the cloze passage. There may not be enough information available to the reader to allow the production of miscues that keep the basic relations in the passage consistent to what was actually written by the author.

The findings of this study are not completely consistent with findings by Rankin (1959). The cloze test used in the present study was a fifth (nth) word deletion. Rankin found that nth word deletion cloze tests were highly related to a test of inferential comprehension. He concluded that an individual must rely more on understanding the interrelationships between ideas in order to complete an nth word deletion cloze test. His conclusion does not explain why, in the present study, the ability to use structural language cues in a cloze test predicted literal comprehension but not inferential comprehension. It is possible that the ability of an individual to utilize syntactic language cues in a cloze test may play a more important role in the analysis of literal reading comprehension skill than previously recognized. The nth word deletion cloze test may be the means through which diagnostic data related to literal comprehension can be obtained.

References

Bormuth, J.R. Validities of grammatical and semantic classifications of cloze test scores. In J.A. Figurel (Ed.), *International Reading Association Conference Proceedings* (Vol. 10). Newark, Delaware: International Reading Association, 1965.

Bormuth, J. R. Factor validity of cloze tests as a measure of reading comprehension. *Reading Research Quarterly*, 1969, *4*, 358-64.

Brown, E. Cloze symposium. In G.B. Schick & M.M. May (Eds.), *Seventeenth Yearbook of the National Reading Conference.* Milwaukee, Wisconsin: The National Reading Conference, Inc., 1968.

Burmeister, L. E. *Reading strategies for secondary school teachers.* Reading, Massachusetts: Addison-Wesley Publishing Company, 1974.

Fry, E.A., A readability formula that saves time. In G.B. Schick & M.M. May (Eds.), *Seventeenth Yearbook of the National Reading Conference.* Milwaukee, Wisconsin: The National Reading Conference, Inc., 1968.

Goodman, K.S. Analysis of oral reading miscues: Applied psycholinguistics. *Reading Research Quarterly,* 1969, *5,* 9-30.

Goodman, K.S. The reading process: Theory and practice. In R.E. Hodges & E.H. Rudorf (Eds.), *Language and learning to read.* Boston, Massachusetts: Houghton Mifflin Co., 1972.

Goodman, K.S. Comprehension centered reading. In E.E. Ekwall (Ed.), *Psychological factors in the teaching of reading.* Columbus, Ohio: Charles E. Merrill Publishing Co., 1973.

Goodman, K. S. Linguistically sound research in reading. In R. Farr, S. Weintraub, & B. Tone (Eds.), *Improving reading research.* Newark, Delaware: International Reading Association, 1976a.

Goodman, K.S. The Goodman taxonomy of reading miscues. In P.D. Allen & D.J. Watson (Eds.), *Findings of research in miscue analysis: Classroom implications.* Urbana, Illinois: National Council of Teachers of English, 1976b.

Goodman, Y. & Burke, C. *Reading miscue inventory.* New York: Macmillan Co., 1972.

Greene, F.P. Modification of the cloze procedure and changes in reading test performances. *Journal of Educational Measurement,* 1965, *2,* 213-217.

Horton, R.J. The construct validity of the cloze procedure: An exploratory factor analysis of cloze, paragraph reading, and structure-of-intellect tests. *Reading Research Quarterly,* 1975, *10,* 248-57.

Karlson, B., Madden, R. & Gardner, E. F. *Stanford Diagnostic Reading Test,* Level II, Form W. New York: Harcourt, Brace and Jovanovich, 1966.

Louthan, V. Some systematic grammatical deletions and their effects on reading comprehension. *English Journal,* 1965, *54,* 295-299.

Ohnmacht, F. & Fleming, J. *Cloze and closure: A second analysis.* (ERIC Document Reproduction Service No. ED 058015), 1972.

Ohnmacht, F., Weaver, W. & Kohler, P. *Cloze and closure: A factoral study.* (ERIC Document Reproduction Service No. EJ 017975), 1970.

Page, W.D. Are we beginning to understand oral reading? *Reading World,* 1974, *13,* 161-70.

Rankin, E.F. The cloze procedure — its validity and utility. In. O.S. Causey & W. Eller (Eds.), *The Eighth Yearbook of the National Reading Conference.* Fort Worth, Texas: The Texas Christian University Press, 1959.

Rousch, P.D. Testing. In P.D. Allen & D.J. Watson, (Eds.), *Findings of research in miscue analysis: Classroom implications.* Urbana, Illinois: National Council of Teachers of English, 1976.

Vaughn, J.L., Tierney R. & Alpert, M. Psycholinguistic analysis of cloze responses. In P.D. Pearson (Ed.), *The Twenty-sixth Yearbook of the National Reading Conference.* Clemson, South Carolina: The National Reading Conference, Inc. 1977.

Weaver, W.W. Theoretical aspects of the cloze procedure. In E.L. Thurston & L.E. Hafner (Eds.), *Fourteenth Yearbook of the National Reading Conference.* Milwaukee, Wisconsin: The National Reading Conference, Inc., 1965.

Weaver, W. & Kingston, A. A factor analysis of the cloze procedure and other measures of reading and language ability. *Journal of Communication,* 1962, *13,* 252-61.

CHARLES W. PETERS
Oakland Schools
Pontiac, Michigan

Assessing reading performance at the secondary level through the utilization of a cognitive self-rating scale

While research has investigated the relationship between self-perception and general academic performance, there is a paucity of research exploring the relationship between one's self-perception and his/her reading ability. Research (Peters & Peters, 1976; Peters, 1977) investigating this topic has revealed that when the score obtained from a cognitive self-rating scale is combined with the score from a general measure of reading performance, the ability to predict reading performance was improved when compared to a method which utilizes a general measure of reading performance.

Since results may vary with the type of comprehension probe utilized (e.g., general vs. specific reading performance, Artley, 1944; Maney, 1958; Peters, Peters & Kaufman, 1975; Sochor, 1958) this study sought to ascertain whether a Cognitive Self-Rating Scale (SRS) would be a valid and reliable measure of reading performance when compared with the results obtained from an instrument designed to measure specific reading skills associated with one cognate area. Therefore, the purpose of this study was to investigate the following questions: (1) Does the SRS measure reading ability? (2) If the SRS is a measure of reading ability, does it distinguish between various levels of reading ability (good and poor)? (3) Is the SRS capable of measuring distinct dimensions of the reading process? (4) Does the SRS predict reading performance?

Method

Subjects

The sample was comprised of 59 students from a suburban junior high school located in an upper middle class community. The students were enrolled in a regular eighth grade English class.

Reading ability was determined on the basis of individual performance on a Content Specific Reading Test (CSRT). Good readers were defined as those students who had a range in raw scores from 35 to 50 while the poor readers were defined as those students who had a range in raw scores from 10 to 25. Students whose raw scores ranged from one to nine were eliminated, because they were considered to be disabled readers who lacked sufficient skills to read correctly the items on the SRS. Those students whose raw scores ranged from 26 to 34 were eliminated to prevent a potential overlap between the good and poor reader categories. There was a possible raw score of 50 on the CSRT. Reliability (.92) and content validity for the CSRT were established in a previous study (Peters, Peters, & Kaufman, 1975).

Materials

If the SRS is to be predicated upon a measureable construct, then it must be comprised of those factors research indicates are related to the reading process. Based upon this premise, the SRS is composed of four factors research has shown are related to the comprehension process: knowledge of concepts, semantic structure of material, cognitive skills associated with the comprehension process, and time required to process printed information. Statements reflecting these factors were incorporated into the 15 item SRS (see Table 1).

Table 1

Cognitive Self-Rating Scale

1. When compared to other members of the class, I would rate myself as a very good reader.
2. I feel that most of the material we read for school is too easy.
3. It takes me a great deal of time to read most of my assignments, because I am a slow reader.
4. I do not read most of the material assigned, because it is too hard.
5. If I selected the material for my classes, I would pick books that are easier to understand.
6. When I complete my reading, I have no difficulty answering the questions asked by my teachers.
7. Reading school materials is difficult for me.
8. When compared to other members of the class, I would rate myself as a poor reader.
9. Most of the material I read in school contains many words that are difficult to understand.
10. I have very little trouble understanding what I read in my classes at school.
11. Material I read in school is easy.
12. I usually do not understand what is happening in the material I read for school.
13. When I read material for school, there are very few words I do not understand.
14. When compared to other members of the class, I would rate myself as an average reader.
15. I have to read material assigned in school over and over to understand it.

The determination of reading ability was based upon students' responses to a four point Likert scale. A value of one was assigned to each response considered to be indicative of good reading behavior. Good readers were defined as those students whose scores ranged from zero to seven. An internal consistency coefficient of .76 for the total scale was determined by the K-R 21 formula.

Design

Five statistical procedures were utilized to analyze the data. First, two correlational matrices were utilized to analyze the relationship between the CSRT and SRS and the interrelationship between items on the SRS. Second, a factor analysis (varimax rotated) was utilized to ascertain whether the SRS was comprised of distinct elements. Third, a t test was utilized to ascertain whether the SRS discriminated between good and poor readers. Fourth, a regression analysis was used to determine the predictive validity of the SRS, and fifth, a crosstabulation of the results from the CSRT and the SRS were performed to ascertain the accuracy of placement by reading ability.

Results

To determine whether the SRS is a measure of reading behavior, it was correlated with CSRT. The results of that comparison reveal that the SRS does measure some dimension of the reading process. The correlation between the two instruments is moderately low ($r = .352$). While the SRS and the CSRT are related, they are measuring significantly different ($p < .01$) dimensions of the reading process.

Since the SRS is a measure of reading behavior, did it measure distinct elements within the reading process? A factor analysis reveals that the SRS contained five factors with Eigenvalues above 1.0. These factors account for 67 percent of the variance. To facilitate the identification of the items in each of the five groups a second factor analysis using the varimax rotation process was performed. The items grouped in each of these five categories were then compared to the original groupings. The results of this comparison reveal that there is no consistent match between the items grouped in the four original categories and the items in the five categories identified in the factor analysis. While the items within the original four categories do not seem to be distinctive groups, the SRS does appear to be comprised of five separate factors. The stability of these groupings is confirmed when the intercorrelational coefficients are examined. All items in each of the five groups were moderately correlated.

The analysis of the results from the t test indicate that the SRS does discriminate between good and poor readers (see Table 2). The difference is significant, $t\,(57) = -2.59$, $p < .01$, means: 20.72 and 16.87. However, when the results are analyzed on a subtest basis (contextual analysis, main ideas, details, sequence, and critical reading), only two subtests discriminate significantly between good and poor readers, main idea, $t\,(57) = -2.21$ $p < .05$, and sequence, $t(57) = -2.05$, $p < .03$.

While the SRS is capable of discriminating between good and poor readers, the regression analysis of pre-CSRT scores and post-SRS scores reveals that the SRS is not a good predictor of general reading performance, $F\,(5, 53) = 2.05$, $p < .052$, since the total variance accounted for was approximately 16 percent. However, when the crosstabulation of placement by reading ability on the SRS and the CSRT is examined, it reveals that the two instruments are extremely consistent in their placement of poor readers. Only 1.7 percent of the students classified as poor readers by the CSRT were classified as good

Table 2

Means of Good and Poor Readers on the CSRT

Subtest	Ability Group	Mean	S.D.
Contextual Analysis	Poor	3.25	1.38
	Good	3.94	1.33
Main Idea	Poor	4.25	1.29
	Good	5.07	0.93
Details	Poor	3.25	1.75
	Good	4.11	1.21
Sequence	Poor	2.0	0.92
	Good	2.88	1.16
Critical Reading	Poor	4.12	1.55
	Good	4.70	1.23
Total Score	Poor	16.87	4.01
	Good	20.72	3.89

readers by the SRS. The converse is true of good readers since 42.4 percent of those students classified as poor by the CSRT were classified as good readers by the SRS.

Conclusions and Implications

The purpose of this study was to determine the utility of an instrument designed to measure students' self-perception of their reading ability. As the results of the correlation between the SRS and the CSRT indicate, the SRS is related to reading ability. This finding is consistent with previous research (Peters, 1977), but more importantly, the SRS is both a measure of general and specific reading ability.

As the results presented in Table 2 reveal, the SRS also has the capability of discriminating between good and poor readers. While the SRS does have this discriminative ability, the percentage of placement of good and poor readers varies from instrument to instrument. The largest discrepancy exists between those readers classified as poor by the CSRT and good by the SRS; forty-two percent of the readers fall into this category. This discrepancy in placement can be partially explained by the fact that the CSRT is too inclusive in its placement of poor readers. It classified 54 percent of the sample as poor readers. Given the educational composition of the population, it is highly unlikely that such a large percentage of the group would be poor readers. Restructuring the parameters used to define poor readers might produce different results.

However, there is a high degree of consistency in the placement of poor readers by both instruments. Only 1.7 percent of the students classified as poor readers by the CSRT were classified as good readers by the SRS. When the total number of poor readers placed by the SRS, 12 percent, is compared with the percentage of poor readers placed by the CSRT, the SRS figure seems more realistic when one considers the average distribution of poor readers in a representative sample.

Since the SRS appears to be an accurate predictor of poor readers, one possible function of the instrument might be as an initial screening device. For example, content teachers who are confronted with large numbers of students might find the SRS highly useful. It could be administered at the beginning of the year or semester. Since the instrument is economical in terms of time required to administer and score it, teachers could quickly identify those students who may be potential poor readers.

The utilization of the cognitive self-rating scale helps illustrate the significance of student input into the diagnostic process. The results of this study suggest that supplemental forms of cognitive assessment can be employed when attempting to measure reading performance. The diagnostic process at the secondary level needs to become less myopic and more comprehensive, because many times critical instructional decisions are made on the basis of one instrument. While more research needs to be done with cognitive self-rating scales, they do have the potential for becoming a very useful instrument.

References

Artley, S. A study of certain relationships existing between general reading comprehension and reading comprehension in a specific subject-matter area. *Journal of Educational Research*, 1944, *37* 464-473.

Maney, E. Literal and critical reading in science. *Journal of Experimental Education*, 1958, *27*, 57-64.

Peters, C., Peters, N. & Kaufman, D. A comparative analysis of reading comprehension in four content areas. In G. McNinch & W. Miller (Eds.), *Reading: Convention and inquiry*. Twenty-fourth Yearbook of the National Reading Conference, 1975, 200-206.

Peters, C. & Peters, N. A systematic approach to predicting reading performance at the secondary level. In W. D. Miller & G. H. McNinch (Eds.), *Reflections and investigations on reading*. Twenty-fifth Yearbook of the National Reading Conference, 1976, 121-128.

Peters, C. Predicting reading performance at the secondary level through the utilization of a student self-assessment inventory. In P. D. Pearson (Ed.), *Reading: Theory, research and practice*. Twenty-sixth Yearbook of the National Reading Conference, 1977, 131-135.

Sochor, E. Literal and critical reading in social studies. *Journal of Experimental Education*, 1958, *27*, 49-56.

JOSEPH J. BLEAKLEY
E. PETER JOHNSEN
University of Kansas

The comprehension of prose: problems in measuring learning outcomes

A major function of our schools is to communicate knowledge. Undoubtedly, the largest proportion of this data is communicated through printed verbal material; the student is required to read in order to learn. Because of its importance, reading comprehension has been widely studied. The educational psychologist has contributed to this burgeoning literature, usually in an effort to find means of facilitating reading comprehension as well as measuring it. Many of these prose learning studies have focused on manipulating the learner's task instructions in order to assess how these instructions influence what is learned. These are certainly commendable efforts. Finding ways to facilitate learning, in this case prose learning, cannot be faulted. However, a careful review of a large sample of the literature reveals that most often the learning outcome that is assessed is educationally trivial. For example, a learner may be asked to recognize or recall a series of words taken from the instructional material. Measuring this type of learning outcome (e.g., names, dates, quantities, technical terms) can only guarantee that rote learning has occurred. Recognizing the surface structure of a linguistic input does not insure that understanding has occurred. Education and, therefore, research should be more interested in meaning than in memory for the verbatim input.

This is not a new cry, of course. Others have made the same plea to measure comprehension rather than memorization (Allen, 1970; Andre & Sola, 1976; Cunningham, 1972; Duell, 1974; Felker & Dapra, 1975; Hiller, 1974; Shavelson, Berliner, Ravitch, & Loeding, 1974; Watts & Anderson, 1971). This paper is an attempt to review research which demonstrates the distinction between rote and meaningful learning of written text and the implications of that distinction for the measurement of learning outcomes. In most reading the retention of discrete and specific information is less important than the meaningful comprehension of basic propositions, the making of required or invited inferences, or the integration of ideas into a meaningful, unified whole. Research has indicated that a learner's memory is quite often an abstract construction created from the text itself rather than a reproduction of it. Since most classroom learning is designed to become incorporated into seemingly stable long-term memory, the distinction between reproductive and constructive processes has implications for the measurement and evaluation procedures of students' learning through reading. A sample of current research findings will, accordingly, be analyzed in terms of the assessment techniques utilized, and the type of dependent variable under scrutiny. A taxonomy of these variables will be presented based upon the distinction between reproductive and constructive processes mentioned above. The validity of assessment devices will be discussed and implications for research design will be made.

Meaningful Learning

In meaningful learning the effort is to acquire understanding. Understanding or comprehension is the result of relating new material to what one already knows (Ausubel, 1968; Haviland & Clark, 1974; Kintsch, 1974; Smith, 1975). If the learner is successful in relating new material to cognitive structure, he has acquired *meaning*. Meanings are more abstract than the words used to convey them. They refer to the substance or "gist" of the message and are generally not word specific. Meaning is a function of those concepts that are used to express it (Kintsch, 1974). A concept can usually be expressed by several synonymous words which still communicate the same meaning. In acquiring meaning, then, readers can store meaning in the form which is most readily incorporated into their particular cognitive structure. They can use those concepts and interrelationships among concepts which are most "meaningful" to them.

If learners are asked to store the actual linguistic input through which the meaning came (a verbatim memorization task), their job is more difficult. Verbatim learning is seemingly less natural than meaningful learning (e.g., Bransford & Franks, 1971), it requires more time (Jones & English, 1926), it does not insure meaningful learning (Burger & Perfetti, 1977; Jones & English, 1926), and it does not persist as well over time (Bobrow, 1970; Sachs, 1967). Educationally, it has been argued by Gagné (Gagné, 1976, Gagné & Briggs, 1974) that it is generalizable *intellectual skills* which make the learner competent, not his store of discrete, factual verbal information. These intellectual skills are understandings or meanings about concepts and rules, for example, which are *usable* by the learner in a variety of situations. Factual knowledge is certainly educationally important but acquiring a system of understandings (i.e., meaningful learning) has been argued by many to be a learning outcome of higher priority.

Memory for Text and Learning Outcomes

Memory for a reading event is *constructive* (e.g., Bransford, Barclay, & Franks, 1972; Cofer, 1973; Neisser, 1967). That is, readers don't interpret the semantic content of a message and store it as an isolated, independent event. Rather, the message content acts as a cue which readers can use to construct a semantic schema (i.e., representation) of the information. These constructed descriptions often contain more information than is represented by the linguistic input (inferences, for example) and, therefore, a purely linguistic analysis will not characterize all of the information available to the reader; prose memory goes beyond and enriches the semantic content presented. Supporting evidence comes from several sources. There is evidence for the recall of inferred meaning which fills gaps in the semantic continuity of the text (Kintsch, Kozminsky, Sterby, McKoon, & Keenan, 1975) and for the integration of the linguistic input into complex, abstracted knowledge structures (Bransford & Franks, 1971; Tzeng, 1975). Individuals also construct memory structures through editing, summarizing, generalizing, etc., in order to minimize information overload (Dawes, 1966; Frederiksen, 1975a; Gomulicki, 1956) and infer and store intended rather than literal meaning when required (Clark & Lucy, 1975; Stillings, 1975).

The ability to reproduce, either by recognition or recall, discrete, factual data corresponds to the lowest level of Bloom's (1956) taxonomy of cognitive objectives. If other, higher level outcomes are available and if memory is constructive such that it builds upon elaborated meanings which go beyond the expressed text, then the measurement of reading outcomes should express these phenomena. If it does not, this type of measurement, at best, ignores much of a reader's memory schema for what has been read and emphasizes trivial outcomes. Knowing the outcomes expected may force the reader to encode specifics at the expense of broader understandings. It may even *misrepresent* what the reader has acquired. Learning and memory expectancies are known to affect a reader's knowledge of a text (Gauld & Stephenson, 1967) and may even hide some of his knowledge of what has been read (Brockway, Chmielewski, & Cofer, 1974).

The point being made is that the effects of a reading experience go beyond what is measured by accurate recall and recognition of factual data. The memory schema for a reading event contains much more than a replica of the stimulus material. Educationally relevant prose learning measurement, therefore, ought to reflect outcomes other than those at the lowest taxonomic level (Bloom, 1956).

A Taxonomy of Reading Outcomes

Since a reader's memory schema is a richly elaborated storehouse of information, it contains more than the facts that were expressed on the written page. Most prose learning studies, unfortunately, use factual verbatim information as the dependent variable of interest. It will be shown that other outcomes that are educationally important and reflect the constructed, elaborated and even productive nature of reading memory have been used only too rarely. Certainly, though, these other "higher level" outcomes are evident in the literature. Our review of the kinds of dependent variables examined in current research along with current understanding of memory for text leads to a proposed logical system or taxonomy of these reading outcomes.

This taxonomy categorizes four types of outcomes or classes of dependent variables: memorization, translation, inference, application. These categories describe behaviors readers may be expected to exhibit after a reading event. They seem to represent dimensions of memory schemata which reflect the characteristics of information from prose which can be acquired, retained, and retrieved by a reader.

Memorization refers to the verbatim remembering of the text or portions of it. *Translation* is recognizing or recalling information (i.e., meaning) in a form different from the manner in which it was expressed in the text. An *inference* is information contained in a memory schema which goes beyond and, therefore, elaborates the expressed text. *Application* is the application of the information in the schema to new events; it requires the reader to transform the information in order to respond to new demands in the environment.

The ordering of the taxonomic levels implies two different hierarchies. One hierarchy is grounded in a series of progressively more advanced levels of educational outcomes or goals. For example, memorization refers to the retention of information, translation and inference to its comprehension, and application to its transferability. The other hierarchy suggests a progression of prerequisite skills. For example, the application of information requires its comprehension as a prerequisite. In the following sections, the levels of the taxonomy will be described in greater detail and exemplified by citations to current research studies.

Level 1: Memorization

Memorization emphasizes the retention and retrieval of specifically expressed information from the text. The name given to this level is not completely satisfactory because of its association in the general population with the intentional act of remembering a text verbatim through repeated drill. The reader is expected to retrieve verbatim data in the form in which it was expressed in the text. If the linguistic form of the information changes (e.g., "10" becomes "ten" or "She generated various ideas" becomes "Many ideas were produced by her"), it has undergone translation and, therefore, belongs to another category of the taxonomy.

Retrieval of this type of information can be through either recognition or recall. Boker (1974) has assessed recognition of specific facts and Mayer (1975) has likewise assessed larger units (e.g., definitions). Recall has been used to measure retention of specific facts with fill-in questions (Frase & Schwartz, 1975), free recall of facts (Shimmerlik & Nolan, 1976), prompted recall of propositions (Ross & DiVesta, 1976) and free recall of entire passages (Jones & English, 1926). Typically, studies have measured the retention of facts such as names, dates, technical terms and quantities (e.g., Frase, 1975; Kaplan, 1974; Rothkopf & Bisbicos, 1967). Specific, factual information is, of course, important and does serve a valuable role in education. It should not be the outcome emphasized, however (Ausubel, 1968).

Level 2: Translation

Translation is synonymous with paraphrase. Acquisition and retrievel of paraphrases of text is considered by some to be required for understanding (Ausubel, 1968). Acquisition of a paraphrase can be best described as "meaning" that has been acquired which can take any number of linguistic forms (Kintsch, 1974). Translation or paraphrase questions are considered by Anderson (1972) and Bloom (1956) to be techniques for measuring understanding. Translation is changing a communication into parallel forms. Paraphrasing a sentence or summarizing a paragraph in one's own words are examples of creating parallel forms. However, they do not even need to be similar in content form. A sentence (semantic) can be represented by a picture (figural), a set of directions by a map, etc. The major advantage of measuring translation outcomes is their sensitivity to comprehension rather than memorization (Bloom, 1956). The emphasis is on meaningful verbal learning rather than rote learning (Ausubel, 1968). These outcomes are usually assessed by having the reader use his own words to attempt to express the "gist" of the text (e.g., Johnson, 1974; Meyer & McConkie, 1973).

Level 3: Inferences

As stated earlier, inferences make a significant contribution to a reader's memory schema for a passage. This is known to be a developmental phenomenon which occurs early but becomes indistinguishable from adults' inferential behavior at around eleven or twelve years of age (Paris & Lindauer, 1976). Many types of inferences are possible as well as necessary. Interpolation and extrapolation of information is one type (cf. Bloom, 1956). Inferring an author's point of view, assumptions, mood, etc., are other examples. These kinds of reading outcomes are, of course, measured in some reading tests (cf. Davis & Davis, 1962).

Requiring readers to make inferences has been shown to influence those readers' retention of related material. In Rickards (1976) asking readers to infer that southern Mala is a desert increased their retention of the ideas in the paragraph over those who were not asked to make the inference. The results would be expected to be the same for readers who spontaneously made the "desert" inference. There is data using other material which has supported this prediction (Christie & Schumacher, 1975). These results demonstrate that it is more than theoretically interesting to discover that readers spontaneously generate and encode inferences during comprehension (Frederiksen, 1975a). Asking readers to go beyond the information given through adjunct questions (Rothkopf, 1966) or test expectation (Brockway, Chmielewski, & Cofer, 1974) can facilitate text comprehension.

Level 4: Application

Application accounts for the "productive" nature of reading memory. That is, rather than simply being retrieved, information can be retrieved and transformed to produce new outcomes; information can be used in new ways or applied in novel situations. This outcome assesses the reader's ability to solve new problems or explain phenomena using his new knowledge (Sanders, 1966). This can range from using new concepts under novel circumstances (Watts & Anderson, 1971) to solving new mathematics problems (Mayer, 1975).

The Classification of Current Research

Table 1 provides an example of how some current research can be classified using the proposed taxonomy. Most of the studies were published during the period 1974-1976 in the *Journal of Educational Psychology* and are representative of current trends in prose learning measurement. Some earlier research has been included because of frequent citation in the more recent papers. Often it was difficult to discover the nature of the reading outcome(s) measured in a research paper (e.g., Walker, 1974). When this was the case, a best guess was used. Otherwise, when a study assessed more than one type of outcome, it was classified according to the "most advanced" level used (e.g., Felker & Dapra, 1975).

Implications for Research Design

Table 1 graphically demonstrates the disproportionate number of research studies sampled here that measured what might be called lower level outcomes. The higher levels are significant dependent variables which should receive more attention in the literature. Accordingly, future research should evaluate the effectiveness of their independent variables in terms of the more educationally meaningful and theoretically more valuable outcomes.

The taxonomy also should provide a means whereby reseach can be compared. This classification system would allow scholars to compare others' research in terms of their own ongoing work. In this way the dissemination of results and the expansion of our knowledge might be facilitated. Probably the single most important contribution that this taxonomy makes is its highlighting of inferences as significant reading outcomes for prose learning research. There is much current work on inferential phenomena (Frederiksen, 1975b; Kintsch, 1977). Future research which assesses them should carefully determine the nature of the inferences sought and just as carefully develop an instrument to measure them. This has not always been the case (McConkie, Rayner, & Wilson, 1973). Studies should also explicitly report the nature of all dependent variables and instruments used to measure them.

Table 1

Research citations classified according to

the four levels of the proposed taxonomy

LEVEL 1: MEMORIZATION

Arkes, Schumacher, & Gardner, 1976
Bassin & Martin, 1976
Boker, 1974
Christie & Just, 1976
Duchastel & Brown, 1974
Frase, 1968
Frase & Kreitzberg, 1975
Frase & Schwartz, 1975
Gagné & Rothkopf, 1975
Kaplan, 1974
Kaplan, 1976a
Kaplan, 1976b
Kaplan & Rothkopf, 1974
Kaplan & Simmons, 1974
Koran & Koran, 1975
LaPorte & Nath, 1976
LaPorte & Voss, 1975
Rasco, Tennyson, & Boutwell, 1975
Rothkopf, 1966
Rothkopf & Billington, 1974
Rothkopf & Billington, 1975
Rothkopf & Billington, 1975
Rothkopf & Bisbicos, 1967
Ross & DiVesta, 1976
Shimmerlik & Nolan, 1976
Siegel, Lautman & Burkett, 1974
Snowman & Cunningham, 1975
Voss, 1974
Walker, 1974
Yekovich & Kulhavy, 1976

LEVEL 2: TRANSLATION

Allen, 1970
Andre & Sola, 1976
Bower, 1974
Crouse, 1974
Hiller, 1974
Johnson, 1974
Jones & English, 1926
King & Russell, 1966
Meyer & McConkie, 1974
Rickards & August, 1975
Royer & Cable, 1975
Royer & Cable, 1976
Shavelson, Berliner, Ravitch & Loeding, 1974
White & Gagné, 1976

LEVEL 3: INFERENCES

McConkie, Rayner, & Wilson, 1973
Rickards, 1975-6
Rickards, 1976
Rickards & DiVesta, 1974

LEVEL 4: APPLICATION

Duell, 1974
Felker & Dapra, 1975
Mayer, 1975
Watts & Anderson, 1971

A reader does more, much more, than passively store independent linguistic information. The actual linguistic data is only a vehicle by which meaning is conveyed, meaning which includes information that was not expressly communicated. The reader is also a user of his stored meanings; he applies it in new ways. As a result, research should be assessing these outcomes and classifying results in a fashion which facilitates their communicability.

References

Allen, D. I. Some effects of advance organizers and level of question on the learning and retention of written social studies material. *Journal of Educational Psychology*, 1970, *61*, 333-339.

Anderson, R. C. How to construct achievement tests to assess comprehension. *Review of Educational Research*, 1972, *42*, 145-170.

Andre & Sola. Imagery, verbatim, and paraphrased questions, and retention of meaningful sentences. *Journal of Educational Psychology*, 1976, *68*, 661-669.

Arkes, H. R., Schumacher, G. M. & Gardner, E. T. Effects of orienting tasks on the retention of prose material. *Journal of Educational Psychology*, 1976, *68*, 536-545.

Ausubel, D. P. *Educational psychology: A cognitive view.* New York: Holt, Rinehart and Winston, 1968.

Bassin, C. B. & Martin, C. J. Effects of three types of redundancy reduction on comprehension, reading rate, and reading time of English prose. *Journal of Educational Psychlogy*, 1976, *68*, 649-652.

Bloom, B. S., Englehart, M.D., Furst, E. F., Hill, W. H. & Krathwohl D. R. (Eds.). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain.* New York: David McKay Company, Inc., 1956.

Bobrow, S. A. Memory for words in sentences. *Journal of Verbal Learning and Verbal Behavior*, 1970, *9*, 363-372.

Boker, J. Immediate and delayed retention effects of interspersing questions in written instructional passages. *Journal of Educational Psychology*, 1974, *66*, 96-98.

Bower, G. Selective facilitation and interference in retention of prose. *Journal of Educational Psychology*, 1974, *66*, 1-8.

Bransford, J. D., Barclay, J. R. & Franks, J. J. Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology*, 1972, *3*, 193-209.

Bransford, J. D. & Franks, J. J. The abstraction of linguistic ideas. *Cognitive Psychology*, 1971, *2*, 331-350.

Bransford, J. D. & Johnson, M. K. Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 717-726.

Brockway, J., Chmielewski, D. & Cofer, C. N. Remembering prose: Productivity and accuracy constraints in recognition memory. *Journal of Verbal Learning and Verbal Behavior, 1974, 13*, 194-208.

Burger, N. S. & Perfetti, C. A. Reading skill and memory for spoken and written discourse. *Journal of Reading Behavior, 1977, 9*, 7-16.

Christie, D. J. & Schumacker G. M. Advance organizers, age, and the recall of the relevant versus irrelevant thematic information. Paper presented at the annual meeting of the American Educational Research Assoc., Washington, D. C., March 1975.

Christie, J. M. & Just, M. A. Remembering the location and content of sentences in a prose passage. *Journal of Educational Psychology*, 1976, *68*, 702-710.

Clark, H. H. & Lucy, P. Understanding what is meant from what is said. *Journal of Verbal Learning and Verbal Behavior*, 1975, *14*, 56-72.

Cofer, C. N. Constructive processes in memory. *American Scientist*, 1973, *61*, 537-543.

Crouse, J. Acquisition of college course material under conditions of repeated testing. *Journal of Educational Psychology*, 1974, *66*, 367-372.

Cunningham, D. J. The retention of connected discourse: A review. *Review of Educational Research*, 1972, *42*, 47-71.

Dawes, R. M. Memory and distortion of meaningful written material. *British Journal of Psychology*, 1966, *57*, 77-86.

Dooling, D. J. & Lachman, R. Effects of comprehension on retention of prose. *Journal of Experimental Psychology*, 1971, *88*, 216-222.

Duchastez, P. C. & Brown, B. R. Incidental and relevant learning with instructional objectives. *Journal of Educational Psychology*, 1974, *66*, 481-485.

Duell, O. K. Effects of type of objectives, level of test questions, and the judged importance of tested materials upon posttest performance. *Journal of Educational Psychology*, 1974, *66*, 225-232.

Felker, D. B. & Dapra, R. A. Effects of question type and question placement on problem solving ability from prose material. *Journal of Educational Psychology*, 1975, *67*, 380-384.

Frase, L. Effect of question location, pacing, and mode upon retention of prose material. *Journal of Educational Psychology*, 1968, *59*, 244-249.

Frase, L. T. & Kreitzberg, V. S. Effects of topical and indirect learning directions on prose learning. *Journal of Educational Psychology*, 1975, *67*, 320-324.

Frase, L. T. & Schwartz, B. J. Effect of question production and answering on prose recall. *Journal of Educational Psychology*, 1975, *67*, 628-635.

Frederiksen, C. H. Acquisition of semantic information from discourse: Effects of repeated exposures. *Journal of Verbal Learning and Verbal Behavior, 1975, 14*, 158-169. (a)

Frederiksen, C. H. Effects of context-induced processing operations on semantic information acquired from discourse. *Cognitive Psychology*, 1975, *7*, 139-166. (b)

Gagne, E. D. & Rothkopf, E. Z. Text organization and learning goals. *Journal of Educational Psychology*, 1975, *67*, 445-450.

Gagne, R. M. *The conditions of learning* (3rd ed.). New York: Holt, Rinehart and Winston, 1976.

Gagne, R. M. & Briggs, L. J. *Principles of Instructional Design*. New York: Holt, Rinehart and Winston, 1974.

Gauld, A. & Stephenson, G. M. Some experiments relating to Bartlett's theory of remembering. *British Journal of Psychology*, 1967, *58*, 39-49.

Gomulicki, B. R. Recall as an abstractive process. *Acta Psychologica*, 1956, *12*, 77-94.

Haviland, S. E. & Clark, H. H. What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1974, *13*, 512-521.

Hiller, J. H. Learning from prose text: Effects of readability level, inserted question difficulty, and individual differences. *Journal of Educational Psychology*, 1974, *66*, 202-211.

Johnson, R. E. Abstractive processes in the remembering of prose. *Journal of Educational Psychlogy*, 1974, *66*, 772-779.

Jones, M. G. and English, H. B. Notional versus rote memory. *American Journal of Psychology*, 1926, *37*, 602-603.

Kaplan, R. Effects of learning prose with part versus whole presentations of instructional objectives. *Journal of Educational Psychology*, 1974, *66*, 787-792.

Kaplan, R. Effects of experience on subjects' use of directions upon learning from prose. *Journal of Educational Psychology*, 1976, *68*, 217-224.

Kaplan, R. Effects of grouping and response characteristics of instructional objectives when learning from prose. *Journal of Educational Psychology*, 1976, *68*, 424-430.

Kaplan, R. & Rothkopf, E. Z. Instructional objectives as directions to learners: Effect of passage length and amount of objective-relevant content. *Journal of Educational Psychology*, 1974, *66*, 448-456.

Kaplan, R. & Simmons, F. G. Effects of instructional objectives used as orienting stimuli or as summary/review upon prose learning. *Journal of Educational Psychology*, 1974, *66*, 614-622.

King, D. & Russell, G. A comparison of rote and meaningful learning connected material. *Journal of Verbal Learning and Verbal Behavior*, 1966, *5*, 478-483.

Kintsch, W. *The representation of meaning in memory*. Hillsdale, N. J.: Lawrence Erlbaum Associates, Publishers, 1974.

Kintsch, W. *Memory and cognition*. New York: John Wiley & Sons, 1977.

Kintsch, W., Kozminsky, E., Sterby, W. J., McKoon, G. & Keenan, J. M. Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning and Verbal Behavior*, 1975, *14*, 196-214.

Koran, M. L. & Koran, J. J. Jr. Interaction of learner apptitudes with question pacing in learning from prose. *Journal of Educational Psychology*, 1975, *67*, 76-82.

LaPorte, R. E. & Nath, R. Role of goal performance in prose learning. *Journal of Educational Psychology*, 1976, *68*, 260-264.

LaPorte, R. E. & Voss, J. F. Retention of prose materials as a function of postacquisition testing. *Journal of Educational Psychology*, 1975, *67*, 259-266.

Mayer, R. E. Forward transfer of different reading strategies evoked by test-like events in mathematics text. *Journal of Educational Psychology*, 1975, *67*, 165-169.

McConkie, G., Rayner, R. & Wilson, S. Experimental manipulation of reading strategies, *Journal of Educational Psychology*, 1973, *65*, 1-8.

Meyer, B. & McConkie, G. What is recalled after hearing a passage? *Journal of Educational Psychology*, 1973, *65*, 109-117.

Mistler-Lachman, J. L. Depth of comprehension and sentence memory. *Journal of Verbal Learning and Verbal Behavior*, 1974, *13*, 98-106.

Neisser, U. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.

Paris, S. G. & B. K. Lindauer. The role of inferences in children's comprehension and memory for sentences. *Cognitive Psychology*, 1976, *8*, 217-227.

Rasco, R. W., Tennyson, R. D. & Boutwell, R. C. Imagery instructions and drawings in learning prose. *Journal of Educational Psychology*, 1975, *67*, 188-192.

Rickards, J. P. Processing effects of advanced organizers interspersed in text. *Reading Research Quarterly*, 1975-6, *11*, 599-622.

Rickards, J. P. Interaction of position and conceptual level of adjunct questions on immediate and delayed retention of text. *Journal of Educational Psychology*, 1976, *68*, 210-217.

Rickards, J. P. & August, G. J. Generative underlining strategies in prose recall. *Journal of Educational Psychology*, 1975, *67*, 860-865.

Rickards, J. P. & DiVesta, F. J. Type and frequency of questions in processing textual material. *Journal of Educational Psychology*, 1974, *66*, 354-362.

Ross, S. M. & DiVesta, F. J. Oral summary as a review strategy for enhancing recall of textual material. *Journal of Educational Psychology*, 1976, *68*, 689-695.

Rothkopf, E. Z. Learning from written instructive material: An exploration of the control of inspection behavior by test-like events. *American Educational Research Journal*, 1966, *3*, 241-249.

Rothkopf, E. Z. Learning from written instructive materials: An exploration of the control of inspection behavior by test-like events. *American Educational Research Journal*, 1977, *3*, 241-249.

Rothkopf, E. Z. & Billington, J. J. Indirect review and priming through questions. *Journal of Educational Psychology*, 1974, *66*, 669-679.

Rothkopf, E. Z. & Billington, M. J. A two factor model of the effect of goal-descriptive directions on learning from text. *Journal of Educational Psychology*, 1975, *67*, 692-704.

Rothkopf, E. Z. & Billington, M. J. Relevance and similarity of text elements to descriptions of learning goals. *Journal of Educational Psychology*, 1975, *67*, 745-750.

Rothkopf, E. Z. & Bisbicos, E. Selective facilitative effects of interspersed questions on learning from written materials. *Journal of Educational Psychology*, 1967, *58*, 56-61.

Royer, J. M. & Cable, G. W. Facilitated learning in connected discourse. *Journal of Educational Psychology*, 1975, *67*, 116-123.

Royer, J. M. & Cable, G. W. Illustrations, analogies, and facilitative transfer in prose learning. *Journal of Educational Psychology*, 1976, *68*, 205-209.

Sachs, J. S. Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics*, 1967, *2*, 437-442.

Sanders, N. M. *Classroom questions: What kinds?* New York: Harper & Row, 1966.

Shavelson, R. J., Berliner, D. C. Ravitch, M. M. & Loeding, D. Effects of position and type of question on learning from prose material: Interaction of treatments with individual differences. *Journal of Educational Psychology*, 1974, *66*, 40-48.

Shimmerlik, S. M. & Nolan, J. D. Reorganization and the recall of prose. *Journal of Educational Psychology*, 1976, *68*, 779-786.

Siegel, A. I. Lautman, M. R. & Burkett, J. R. Reading grade level adjustment and auditory supplementation as techniques for increasing textual comprehensibility. *Journal of Educational Psychology*, 1974, *66*, 895-902.

Smith, F. *Comprehension and learning*. New York: Holt, Rinehart and Winston, 1975.

Snowman, J. & Cunningham, D. J. A comparison of pictorial and written adjunct aids in learning from text. *Journal of Educational Psychology*, 1975, *67*, 307-311.

Stillings, N. A. Meaning rules and systems of inferences for verbs of transfer and possession. *Journal of Verbal Learning and Verbal Behavior*, 1975, *14*, 453-470.

Tzeng, O. J. L. Sentence memory: Recognition and inferences. *Journal of Experimental Psychology: Human Learning and Memory*, 1975, *1*, 720-726.

Voss, J. F. Acquisition and non-specific transfer effects in prose learning as a function of question form. *Journal of Educational Psychology*, 1974, *66*, 736-740.

Walker, B. Effects of inserted questions on retroactive inhibition in meaningful verbal learning. *Journal of Educational Psychology*, 1974, *66*, 486-490.

Watts, G. H. & Anderson, R. C. Effects of three types of inserted questions on learning from prose. *Journal of Educational Psychology*, 1971, *62*, 387-394.

White, R. T. & Gagné, R. M. Retention of related and unrelated sentences. *Journal of Educational Psychology*, 1976, *68*, 843-852.

Yekovich, F. R. & Kulhavy, R. Structural and contextual effects in the organization of prose. *Journal of Educational Psychology*, 1976, *68*, 626-635.

JOSEPH L. VAUGHAN, JR.
KEITH E. MEREDITH
University of Arizona

Reliability of the cloze procedure as assessments of various language elements

As a measure of reading comprehension, the cloze procedure "... is possibly the most thoroughly validated and sophisticated method of testing presently used in education" (Bormuth, 1975, p. 82). However, investigations of cloze have frequently provided "... only correlational data between cloze test scores and comprehension test scores ... in the absence of direct true score realiability coefficients ..." (Alameida, 1973, pp. 28-29). The purposes of this study were (1) to examine the reliability of the cloze procedure as a measure of semantic awareness in post-reading situations and (2) to examine the reliability of cloze scores as an index of students' syntactic fluency.

Method

The subjects for this investigation were 298 eighth grade students in Tucson, Arizona. Each student read two science-related selections (Passage A about Cancer and Passage B about Superstitutions) determined to be of eighth level reading difficulty as predicted by the Fry Readability Graph (1967). The study was conducted on two consecutive days with the students randomly assigned to two groups. On the first day, one group read Passage A and responded to the cloze test for that passage; the other group read Passage B and responded to the cloze test for that passage. On the second day, the passages were reversed for the groups.

The cloze tests were based on a fifty-item, every-seventh deletion pattern. The procedures for the students were (1) read the passage, (2) turn in the passage and get the cloze test, and (3) respond to the cloze test. No time limit was placed on either the reading or the cloze test. So that scores of syntactic fluency could be derived as well as scores of semantic awareness, students were directed to "Fill in every blank. Leave no blank empty, even if you have a take a wild guess."

Each blank on the cloze tests was examined for exact-replacement, acceptable synonyms, and syntactic acceptability. Synonyms were determined acceptable when agreed to by at least three of four evaluators. In the syntactic acceptability category, a response was considered correct if it agreed with the part of speech *and* the proper use in context of the deleted item.

Based on these categories, the following five scores were derived: (1) Semantically Acceptable, Exact-Replacement (ER); (2) Semantically Acceptable, Exact-Replacement-Plus-Synonyms (ER + SNM); (3) Syntactically Acceptable (STC); (4) Syntactically Acceptable Minus Exact-Replacement (STC − ER); and (5) Syntactically Acceptable Minus Exact-Replacement-Plus- Synonyms (STC − (ER + SNM)). Each of these scores was expressed as a percentage, using the following formulas:

$$\text{ER} = \frac{\text{ER}}{50}$$

$$\text{ER} + \text{SNM} = \frac{\text{ER} + \text{SNM}}{50}$$

$$\text{STC} = \frac{\text{STC}}{50}$$

$$\text{STC} - \text{ER} = \frac{\text{STC} - \text{ER}}{50 - \text{ER}}$$

$$\text{STC} - (\text{ER} + \text{SNM}) = \frac{\text{STC} - (\text{ER} + \text{SNM})}{50 - (\text{ER} + \text{SNM})}$$

All of these scores were examined for internal consistency reliability using Cronbach's Alpha. Parallel form reliability was determined with Pearson product-moment correlations. While specific results will be mentioned in the text, Table 1 provides a synoptic view of all the coefficients.

Results

Coefficients for the two semantic scores (exact-replacement and exact-replacement-plus-synonyms) ranged from .86 to .92 for internal consistency and .79 to .81 for parallel form reliability. The difference between internal consistency coefficients and parallel form coefficients was significant ($p \leqslant .01$) as determined by Feldt's W test (1969) and Fisher's Z test (Glass and Stanley, 1970).

The internal consistency coefficients of the syntactically acceptable score were .95 and .94 for the two selections. The parallel form coefficient was .86. Internal consistency coefficients differed significantly ($p \leqslant .01$) from the parallel form coefficient. All coefficients for the syntactically acceptable score were significantly greater ($p \leqslant .01$) than the coefficients for any of the other types of scores.

For the two syntactic scores derived by extracting part or all of the semantic scores, the internal consistency coefficients ranged from .67 to .81. For each of these two scores, the coefficients for Passage A were significantly greater ($p \leqslant .01$) than the coefficients for Passage B. The parallel form coefficients for these scores were .73 and .67, and these coefficients were significantly lower ($p \leqslant .01$) than their respective internal consistency coefficients.

Discussion

For all practical purposes, the reliability of these five types of cloze scores is clearly established. However, four statistically significant findings deserve examination. They are:

1. For all the scores, the internal consistency coefficients are significantly higher ($p \leqslant .01$) than the parallel form coefficients;
2. The internal consistency coefficients of the two semantic scores are significantly higher ($p \leqslant .01$) than the internal consistency of the corresponding syntactic scores;

Table 1

Reliability Coefficients of

Various Cloze Scores

Internal Consistency Coefficients

	.r Passage A (Cancer)	.r Passage B (Superstitions)
Semantic, Exact-Replacement Only	.86	.89
Semantic, Exact-Replacement-Plus-Synonyms	.90	.92
Syntactic, Total	.95	.94
Syntactic, Exact-Replacement Extracted	.81	.72
Syntactic, Exact-Replacement-Plus-Synonyms Extracted	.77	.67

Parallel Form Coefficients

	r.	corrected .r
Semantic, Exact-Replacement Only	.79	.91
Semantic, Exact-Replacement-Plus-Synonyms	.81	.89
Syntactic, Total	.86	.92
Syntactic, Exact-Replacement Extracted	.73	.96
Syntactic, Exact-Replacement-Plus-Synonyms Extracted	.67	.93

3. The reliability coefficients of the total syntactically acceptable score are significantly greater ($p \leqslant .01$) than the coefficients of each of the other scores; and
4. When corrected for attenuation, the parallel form coefficients of the syntactic scores are significantly higher ($p \leqslant .01$) than the parallel form coefficients of the corresponding semantic scores.

Passage Dependency

The significant differences between the internal consistency coefficients and the parallel form coefficients suggests that all of these close scores may be somewhat passage dependent. For most practical purposes, given the relatively high parallel form coefficients (.67 to .86), the degree to which reliability is passage dependent is likely to be of minimal concern. However, this discrepancy may be of particular interest to a researcher who is concerned with (1) whether this tendency toward passage dependency is representative of the construct underlying cloze as a measure of various language factors and (2) if it is representative, why this might be so. These are issues of validity and cannot be answered by these findings. What these data do suggest is that all of these scores are dependent upon the passage used and a caution must be raised about the generalizability of cloze scores. This is a reasonable caution for the use of any score of reading comprehension since comprehension (however defined) may be specific to a given passage. To generalize students' ability on the basis of performance on a single test of any nature is hazardous at best, and to do so on the basis of a cloze test is no exception.

Differences in Internal Consistency

The significant differences between the internal consistency coefficients of the semantic and corresponding syntactic scores suggest that the semantic scores are more reliable. Such a conclusion, however, is probably not valid due to the derivation of those syntactic scores that are derived by extracting part or all of the semantic scores. In the process of extracting the semantic scores, the variability within the resulting syntactic scores was substantially restricted.

The implications of this restricted variability become clearer with a recognition that these syntactic scores are based on an unequal number of responses since the extraction factor for each individual varies considerably depending upon the number of responses that were semantically acceptable. Thus, the number of responses considered in these syntactic scores is lower than the consistent number (50) of scored semantic responses. This tends to reduce the internal consistency coefficients of these syntactic scores because the potential variability is both inconsistent and less than that potential for the semantic scores.

In effect, then, this seems to be a statistical finding due more to a scoring procedure than a finding with practical implications. Given this computational limitation, it is likely that the internal consistency coefficients for these syntactic scores are deflated, and if some reasonable means could be identified to correct for the limitation, these coefficients would be substantially higher. Unfortunately, there is apparently no way to correct for this problem without significantly increasing the length of the cloze test, and that seems to be unnecessary in light of the relatively high coefficients.

This limitation of restricted variability may become even clearer when examined exclusively within the syntactic scores derived by extracting part or all of the semantic factor. Among the internal consistency coefficients of these two scores, a significant difference exists between Passage A and Passage B. This is not considered an important finding; yet it is worthy of discussion as clarification and as an example of another issue related to cloze scores.

For Passage A the coefficients of internal consistency for these two scores were .81 and .77, but for Passage B they were .72 and .67. Those for Passage A were significantly higher ($p \leqslant .01$) than those for Passage B as determined by Feldt's W test. One might be tempted to surmise that these results support the premise that reliability of various cloze scores is passage dependent. However, this difference becomes clearer when examined in relation to the issue of restricted variability.

The percentage means of the two semantic scores for these passages were respectively 29 and 42 for the exact replacement and 35 and 58 for the exact replacement plus synonyms. Clearly, Passage B was easier for these students than Passage A, irrespective of the determination by Fry's graph that these two selections were of eighth level difficulty. Because the students obtained higher semantic scores on Passage B, when those scores were subtracted to determine the corresponding syntactic scores for Passage B, the remainder was significantly less than the remainder for Passage A. The result was that the variability within these two syntactic scores for Passage B was restricted. This, in turn, deflated the internal consistency coefficients for Passage B. Had this restricted variability not existed, it is likely that these coefficients for the two passages would have been as akin as the corresponding coefficients for the semantic scores. Thus, it would be inappropriate to support the premise of passage dependency with the reliability of these scores.

Syntactically Acceptable Score

An intriguing prospect is offered by the syntactically acceptable score in that it may be the most accurate index of students' comprehension as measured by cloze scores because it does include both syntactic and semantic awareness. What is meant by reading comprehension is an issue to be debated, however, one perspective, typically called "psycholinguistic," suggests that it should include semantic understanding and syntactic fluency. The cloze test, in general, requires that a student utilize understandings about the structure of language and the content of a specific selection to replace items that have been deleted from the selection. Given the exceptionally high reliability coefficients for the total syntactic score and given that those coefficients are significantly higher than all the other scores, it is not an unreasonable hypothesis that the total syntactically acceptable score is a more accurate index of reading ability than other scores as determined by the cloze procedure. It would prove to be a cumbersome procedure to be used by classroom teachers if they were to analyze responses for syntactic acceptability, but that is not too much to ask of researchers if this score were found to be a more accurate index of reading comprehension than the other scores. To ignore syntactically acceptable responses on a cloze test seems tantamount to discarding pertinent information about the students' understanding of the selection, especially if certain psycholinguistic assumptions are accepted. Since implications cannot be extended beyond the existing data, the most that can be suggested is a need for a construct validation of the total syntactic score, along with the others discussed here, to shed more light on what the various scores are revealing about students' reading comprehension.

Parallel Form Coefficients

The significant difference between the parallel form coefficients of the semantic and corresponding syntactic scores may also have implications for the construct validity of these various cloze scores. Initial observation of these coefficients indicates that the coefficients for the semantic scores are significantly greater than those for the syntactic scores. This can best be explained by the deflated internal consistency coefficients caused by restricted variability in the syntactic scores. That the internal consistency coefficients are deflated has a concommitant effect on the parallel form coefficients. Thus, to examine the relationship of the parallel form coefficients of the two scores appropriately, it was necessary to correct the parallel form coefficients for attenuation (Nunnally, 1967) given that the internal consistency coefficients for the two selections were not perfect.

When corrected for attenuation, the parallel form coefficients indicate that the syntactic scores are significantly ($p \leqslant .01$) more reliable than the semantic scores. An explanation for this is that the syntactic scores derived by extracting the semantic variable does indeed assess a syntactic factor. If it does, it is reasonable that the syntactic score would be more reliable across passages (parallel forms) because the structure of the language remains relatively constant while the content of the selections (semantic factor) tends to vary.

Conclusion

All of the cloze scores examined in this study are highly reliable for practical purposes. The significant differences between internal consistency reliability and parallel form reliability lead to one primary implication. Caution must be exercised when using cloze scores based on a single passage to draw conclusions about students' reading ability. It follows, therefore, that multiple cloze selections should be used for accurate assessments of reading ability.

References

Almeida, P. *The cloze test as a measure of reading comprehension.* Unpublished manuscript offered as a qualifying paper. Cambridge, Mass.: Harvard University, 1973.

Bormuth, J. The cloze procedure. In W. D. Page (Ed.), *Help for the reading teacher: New directions in research.* Urbana, Illinois: NCRE/ERIC, 1975, 60-90.

Feldt, L. A test of the hypothesis that Cronbach's alpha or Kuder-Richardson coefficient twenty is the same for two tests. *Psychometrika*, September 1969, 363-373.

Fry, E. A readability formula that saves time. *Journal of Reading, 11*, 513-516.

Glass, G. & Stanley, J. *Statistical methods in education and psychology.* Englewood Cliffs, N. J.: Prentice-Hall, 1970.

Goodman, Y. & Burke, C. *Reading miscue inventory.* New York: MacMillan, 1972.

Nunnally, J. *Psychometric theory.* New York: McGraw-Hill, 1967.

Rankin, E. The cloze procedure revisited. *Interaction: Research and practice in college- adult reading.* Louisville: National Reading Conference, 1974, 1-8.

KEITH MEREDITH
JOSEPH VAUGHAN
University of Arizona

Stability of cloze scores across varying deletion patterns

Developers of cloze instruments have proposed a variety of procedures for selecting words to be deleted from a passage. Taylor, in his introduction of the cloze procedure (1953), depended on the principles of statistical random sampling for selection of words to be deleted. Random deletion is accomplished by using a table of random numbers to delete words up to a predetermined percentage of the words in the passage, e.g., a random 20 percent. More recently various modified cloze procedures have been proposed; however, the predominant number of studies use some approximation of random deletion. Random deletion has typically been approximated by deletion of every *n*th word with the value of *n* specified by the cloze developer.

Every *n*th word deletion patterns have been employed frequently in research studies where true random deletion patterns have been virtually ignored. Two primary reasons may be presented for use of *n*th word deletion patterns rather than random deletion patterns.

1. Developers consider the *n*th word deletion patterns easier to construct than random deletion patterns.

2. Researchers (Fillenbaum, Jones, & Rapport, 1963; MacGinitie, 1961) have concluded that four or more words are needed between blanks in order to obtain the most reliable and presumably valid measures. A conclusion based on these findings has been that items attain their maximum independence from one another when a minimum of four words intervene between each item (Bormuth, 1975). In true random deletion the number of intervening words between items may vary from zero to a large number of words.

Given these advantages one might still question whether a form of systematic sampling such as *n*th word deletion should be used in measurement of a construct which is potentially cyclic in nature and could result in biased samples. Researchers and practitioners alike are interested in a stable as well as reliable measure of a subject's ability to read a passage. That is, different forms constructed on the same passage should not only rank the subjects the same (reliability) but also provide the same index of a subject's ability on that passage (stability). The stability of *n*th word deletion has already been questioned in a study by Bormuth (1964) where he varied the starting point of an every-fifth word deletion pattern to produce all five possible forms for each of 20 passages. Significant differences were found among the means of the various forms. Bormuth concluded that more than one cloze test should be constructed for a given passage when using *n*th word deletion patterns. If one follows Bormuth's recommendation of multiple forms, the first advantage of *n*th word deletions, i.e., ease of construction, is eliminated. One rarely finds multiple forms being used in the literature possibly due to its impracticality. Given that one form and one form only is usually developed over a passage, investigations of the stability of the resulting index are needed comparing *n*th word deletion patterns to random word deletion patterns.

The second advantage of *n*th word deletion patterns, i.e., maximum reliability, has been built on evidence from research on *n*th word deletion patterns only. When patterns other than *n*th word deletion patterns were constructed for comparisons, these patterns were usually of the selective or rational deletion variety. Little direct evidence exists which compares the reliability of cloze tests based on *n*th word deletions to the reliability of cloze tests based on random word deletions. If the conclusion concerning maximum item independence is correct, then one would expect *n*th word deletion patterns to result in more reliable measures than random word deletion patterns.

The purpose of this study was to investigate the stability and reliability of scores resulting from cloze tests based on *n*th word deletion patterns as compared to cloze tests based on random word deletion patterns. Stability was assessed by comparing the variability of means resulting from the different forms which could be developed using a specific deletion pattern. Reliability was assessed using Cronbach's Alpha, an index of internal consistency.

Method

A reading passage was chosen which was estimated to be at the sixth grade reading level as determined by the *Fry Readability Graph* (Fry, 1968). An every-fifth word deletion rate was used for the *n*th word deletion pattern as this rate appears to be the most widely used and accepted deletion rate. All five possible cloze test forms using as every-fifth word deletion pattern were developed over a 250 word segment from the selected passage. Five additional cloze test forms were developed by deleting a random sample of 50 words from the same 250 word segment. The words deleted for each of the random deletion patterns were determined by drawing a sample of 50 numbers from 1 to 250. Words in the segment were deleted which corresponded to these 50 numbers. A minimum of 25 words preceded the first blank and minimum of 25 words followed the last blank of the segment.

The ten cloze test forms, five every-fifth word deletion test forms and five random deletion test forms, were randomly assigned for administration to 446 seventh and eighth grade students at a junior high school in Tucson, Arizona. A sixth grade passage was used as the basis for the tests as standardized tests indicated the school to be somewhat below national average.

The tests were administered without pre-reading. Students were directed to fill in each of the blanks with the word that they thought best belonged in the blank. Students were encouraged to fill in every blank even if they were not sure of what word should go in the blank. Exact word replacement was used in scoring the tests.

Results

Table 1 reports the means, standard deviations, and Cronbach's Alpha for each of the forms for each deletion pattern.

A 2x5x2 (Deletion Pattern x Form x Grade) factorial analysis of variance was used to analyze the differences among the means. A factorial analysis was selected to allow investigation of the grade effects. Pattern by form interactions will have no meaning as there is no basis for particular combinations of forms across deletion patterns.

The passage, as would be expected, was more difficult for seventh graders than eighth graders, $F(1,426) = 12.60, p < .01$. No interaction existed with grade. Therefore the effects of pattern and form are generalizable across grade. The pattern main effect and form main effect are not directly interpretable because of the significant interaction between them, $F(4,426) = 6.55, p < .01$. Tukey

Table 1

Means, Standard Deviations, and Cronbach Alphas for Each Form within Each Deletion Pattern

Deletion Pattern		Forms 1	2	3	4	5
Every-Fifth Word	*M*	24.8	18.8	22.7	17.4	17.3
	SD	7.7	7.8	5.8	7.8	6.7
	Alpha	.86	.90	.80	.89	.84
Random	*M*	17.8	19.6	16.3	19.0	15.9
	SD	6.6	7.0	8.2	9.8	6.9
	Alpha	.85	.85	.90	.93	.84

posthoc procedures were used to compare form means with one another within deletion pattern. Of the ten comparisons possible among forms within the every-fifth word deletion pattern, five were significant (p <.01). None of the ten possible comparisons were significant for random deletion forms even at the .05 level.

The range of Cronbach Alphas was from .80 to .90 for the every-fifth word deletion patterns and from .84 to .93 for the random deletion patterns. The average alpha for fifth-word deletion patterns was .858 while the average for random deletion patterns was .874. One may certainly conclude that the Alphas for every-fifth word deletion tests are not significantly greater than Alphas for random deletion tests.

Table 2

Analysis of Variance Summary Table

Source	df	Sum of Squares	Mean Square	F	Probability
Deletion Pattern	1	653.15	653.15	11.73	.001
Form	4	986.95	246.75	4.43	.002
Grade	1	701.37	701.37	12.60	.001
Pattern x Form	4	1457.77	364.44	6.55	.001
Pattern x Grade	1	10.62	10.62	.19	.663
Form x Grade	4	59.77	14.94	.27	.898
Pattern x Form x x Grade	4	128.74	32.19	.58	.679
Error	426	23712.29	55.66		

Discussion and Conclusions

As in Bormuth's study, significant differences were found among the means from the five every-fifth word deletion cloze tests. Yet, no significant differences were found among means from five random deletion cloze tests. This indicates that random deletion patterns result in a more stable index representing students' abilities on a particular passage. Therefore, if only one cloze test is to be constructed over a passage then a random deletion pattern would be more desirable than an every-fifth word deletion pattern. This is especially true when one considers that no sacrifice in reliability is made.

The range of means for the every-fifth word deletion forms was 7.5 points or a difference of 15 percent. The range for random deletion means was 3.7 points or 7.4 percent. This percentage range would make considerable difference if one is using criterion cutoffs for determining frustrational, instructional, or independent levels for a passage. Again random deletion patterns would be more desirable than every-fifth word deletion patterns. However, new criteria for cutoffs would need to be developed because of the random deletion tests being more difficult than every-fifth word deletion patterns.

REFERENCES

Bormuth, J. R. *Relationships between selected language variables and comprehension ability and difficulty* (Cooperative Research Project No. 2082). Los Angeles: UCLA, 1964.

Bormuth, J. R. Literacy in the classrom. In W. D. Page (Ed.), *Help for the reading teacher, new directions in research.* Urbana, Illinois: National Conference on Research in English, 1975.

Fillenbaum, S., Jones, L., & Rapoport, A. Predictability of words and their grammatical classes as a function of rate of deletion from speech transcript. *Journal of Verbal Learning and Verbal Behavior*, 1963, *2*.

Fry, E. A readability formula that saves time. *Journal of Reading*, 1968, *11*, 513-516, 575-578.

MacGinitie, W. H. *Contextual constraint in English prose.* Ph.D. dissertation, Teachers College, Columbia University, 1960.

Taylor, W. L. Cloze procedure: A tool for measuring readability. *Journalism Quarterly*, 1953, *30*, 415-433.

JOHN M. BRADLEY
KEITH MEREDITH
University of Arizona

The reliability of the Maze procedure for intermediate and junior high school students

Assuming that reading primarily involves the construction of meaning from printed language, the valid assessment of comprehension is crucial for appropriate reading instruction. In spite of an obvious need, there presently appears to be no practical, reliable, and valid procedure available which can be used by a teacher to assess comprehension during instruction.

To be of practical value, a test should have the following attributes: (a) The format should be uncomplicated. (b) The test should be easily administered to groups requiring a reasonable amount of time to give. (c) Scoring should be objective. (d) The test should result in the same outcomes no matter who constructs it. Test reliability must exist both in terms of consistency among items within a test and comparability of data obtained from tests constructed on parallel passages. Validity exists if the test data can be appropriately used for decision making in the instructional setting.

One informal means of assessing reading comprehension which appears to be practical for classroom use is the maze procedure, a modification of the cloze procedure (Guthrie, 1973). The maze procedure utilizes multiple-choice items while the cloze procedure utilizes completion items. Guthrie recomends three options per item: (a) the correct word, (b) a syntactically incorrect word, and (c) a syntactically correct but semantically incorrect word.

The length suggested for a maze test passage is 140 to 160 words, with every fifth to every tenth word replaced with a maze multiple-choice item. According to Guthrie, Seifert, Burnham and Caplan (1974), a maze test can be administered easily to a group in ten to fifteen minutes and allows for objective scoring. In the same paper, they also suggested that teachers can be easily trained to construct valid maze tests.

Some evidence concerning the reliability and validity of the maze procedure is currently available. Gallant (1965) found a maze type test to have high internal consistency and adequate concurrent validity for a sample of first grade students. Guthrie (1973) administered maze tests to 12 primary grade and 24 intermediate grade students and obtained high internal consistency and acceptable concurrent validity. Bradley, Ackerson and Ames (in press) found adequate maze procedure parallel form reliability with a sample of 246 second grade students. Other than limited preliminary evidence offered by Guthrie (1973), no documented evidence is available concerning the reliability and validity of the maze procedure for use with intermediate and junior high school students.

Standards of mastery are also an important aspect of an informal reading test. Guthrie's suggested maze test performance standards for the instructional reading level are 50 to 85 percent accuracy, although no evidence was offered for their validity.

The maze procedure appears to hold promise as an informal reading comprehension test for use to initiate instructional placement and monitor progress. The present study was undertaken for the following purposes: (1) The reliability of maze procedure forms administered to intermediate and junior high school students was studied. It was hoped that the present study would yield maze

195

forms sufficiently reliable to make possible future research into maze procedure validity. (2) Placement decisions resulting from maze procedures were compared to placement decisions resulting from cloze procedures on common passages. The cloze procedure was selected as a measure for comparison with maze because cloze instructional placement standards have been empirically derived (Bormuth, 1967, 1968; Ransom, 1968) and the cloze procedure appears to provide a valid measure of reading comprehension (Bormuth, 1963, 1969a, 1969b; Gallant, 1965; Jenkinsen, 1957; Rankin, 1957; Taylor, 1956, 1957). (3) An item analysis was conducted to investigate the adequacy of Guthrie's recommended maze test option types. (4) The consistency of cloze scores and maze scores across parallel passages was determined.

Method

Materials

Six reading passages, each consisting of at least 300 words of running text, were selected for use to construct the maze and cloze tests. Two different passages (Passage A and Passage B) were selected for use at each of three grade levels (fourth, sixth, and eighth). The passages used to construct the fourth and sixth grade tests were taken from the fourth and sixth readers of the *Houghton Mifflin Reading Program* (Durr, Windley & McCourt, 1971). The two passages selected for use at eighth grade were taken from appropriate juvenile reading sources. The *Fry Readability Graph* (Fry, 1968) was also used to aid in the selection process.

The following procedures were employed separately for each of the six test passages. The first and last sentences in each passage were left intact. The first and last sentences in each passage were left intact. The words in each passage identified for use as items were selected using an every seventh-word-deletion pattern. The first word to be deleted was selected using random procedures. A maze test was constructed by replacing each identified word with a maze multiple-choice item. A cloze test was constructed by deleting each identified word and replacing it with a blank of fifteen spaces. This procedure resulted in two 40-item maze forms (Maze A and Maze B) and two 40-item cloze forms (Cloze A and Cloze B) prepared for each of three grade levels (fourth-, sixth-, and eighth-grade). Exact word scoring was used with the cloze test.

Subjects

The sample consisted of 335 fourth, sixth, and eighth grade students attending four schools in Tucson, Arizona. The sample consisted almost entirely of Anglo-suburban dwelling students from middle income families.

Procedure

During May, 1977, the maze and cloze tests were administered to the sample. Intact classes were randomly assigned to either procedure. A short training session was provided to familiarize the students with the purpose and format of the testing procedure (either maze or cloze). Each student received two forms to complete, one form for each test passage (Form A and Form B). Half the students first completed the test forms for Passage A while the other half first completed those for Passage B. The particular form (A or B) given first to a student was determined randomly. The students in the maze groups were given directions consistent with those suggested by Guthrie, Seifert, Burnham and Caplan (1974). Those in the cloze groups were directed to, "Fill in each cloze blank with the word that makes the best sense in terms of the story."

Results and Discussion

Table 1 reports numbers of subjects, means, standard deviations, and Cronbach Alphas for each form of maze and cloze for each grade level.

Table 1

Means, Standard Deviations and Alpha Coefficients
For Cloze and Maze Passages

Grade Level and Form	*n* Cloze	Maze	Mean Cloze	Maze	*SD* Cloze	Maze	Alpha Cloze	Maze
Fourth Grade								
Form A	56	59	17.7 (44%)	37.4 (94%)	4.8	2.6	.76	.71
Form B	56	59	15.0 (38%)	33.1 (83%)	4.6	3.7	.75	.73
Sixth Grade								
Form A	57	48	10.2 (26%)	34.8 (87%)	3.5	3.7	.65	.74
Form B	57	48	12.6 (32%)	35.6 (89%)	4.4	4.6	.74	.85
Eighth Grade								
Form A	49	66	16.1 (40%)	36.3 (91%)	5.5	3.8	.80	.80
Form B	49	66	18.5 (46%)	38.5 (96%)	6.8	1.8	.87	.62

Constructing maze tests using Guthrie's recommendations resulted in tests which were considerably easier than corresponding cloze tests. Every maze test had a ceiling effect where scores were approaching the upper limit of the score range. Contrasted to this are the cloze passage means which were at the opposite end of the difficulty continuum.

Maze and cloze scores converted to percentages were interpreted in terms of available functional reading level criteria for the maze procedure (Guthrie et. al., 1974) and the cloze procedure (Bormuth, 1967; Rankin & Culhane, 1969).

The maze mean scores on five of the six forms ranged from 87 to 96 percent, identifying students at the independent reading level. The one exception was the mean score obtained with Maze Form B at the fourth grade level. That mean score (83 percent) fell near the upper limits of the instructional level.

Different conclusions concerning student placement were generated by the cloze procedure results. Cloze forms at the fourth and eighth grade levels placed students at the instructional level (38 to 46 percent) while cloze forms at the sixth grade level placed students at the frustrational level (26 to 32 percent).

The ease of the maze task resulted in restricted variability in Form A, grade four and Form B, grade eight. Variability was most restricted in the cloze scores at the sixth grade level because of the difficulty of the task. The alpha coefficients were found to be directly related to the ceiling effect obtained with the maze tests and to the floor effect obtained on the sixth grade cloze tests. To the degree that variability was restricted, internal consistency was lowered. The two lowest indices within the maze tests (Maze Form A, fourth grade and Maze Form B, eighth grade) were where the ceiling effect was most pronounced. The two lowest indices within cloze tests were at sixth grade where the floor effect was most pronounced and variability was most restricted.

As indicated by the high maze test means, maze procedure items were very easy. Therefore the test did not discriminate among subjects. Options must be redesigned for these age levels in order for the maze procedure to become a viable testing procedure. Data from this study provided two clues for option redesign.

1. Incorrect options were not attractive as would be required for an adequate testing procedure. Forty to sixty percent of the subjects should respond to incorrect options for an item to have maximum efficiency as a discriminator (Thorndike, 1971). Recommended maze option distractors did not meet this criterion as an average of only 17% of the responses were drawn to the distractors.
2. The syntactically correct, semantically incorrect option was found to be a more effective distractor than the syntactically incorrect option. Approximately two-thirds of the incorrect responses were of the syntactically correct, semantically incorrect type. Parallel form reliability coefficients were computed for grades four, six, and eight for cloze (.65, .51, and .70 respectively) and maze (.58, .59, and .50 respectively). These reliability coefficients were found to be directly related to internal consistency coefficients which, in turn, were related to restricted variability. Differences in parallel form reliability coefficients can be, to a large extent, explained by the restricted score range.

Conclusions and Recommendations

The maze procedure as proposed by Guthrie appears to be inappropriate for use with intermediate and junior high school students. Maze procedures resulted in student placement predominatly at the independent level while cloze procedures constructed on the same passages resulted in student placement at the instructional or frustration level. The use of the maze procedure produced a ceiling effect which reduced score variability, parallel form reliability, and therefore the test's ability to discriminate.

It is possible that with certain modifications, the maze procedure could be made appropriate for use with intermediate and junior high school students. Possible maze procedure modifications might include: (a) discarding the option type utilizing a syntactically incorrect word, (b) devising new option types, e.g., semantically correct within sentence but semantically incorrect within passage, (c) increasing the number of options per item.

References

Bormuth, J. Cloze as a measure of readability. In Figurel, J. A. (Ed.), *Reading as an intellectual activity: Proceedings of International Reading Association*, 1963, *8*, 131-134.

Bormuth, J. Readability: A new approach. *Reading Research Quarterly*, 1966, *1*, 79-135.

Bormuth, J. Comparable cloze and multiple-choice comprehension test scores. *Journal of Reading*, 1967, *10*, 291-299.

Bormuth, J. Cloze test reliability: Criterion reference scores. *Journal of Educational Measurement*, 1968, *5*, 189-196.

Bormuth, J. Factor validity of cloze tests as measures of reading comprehension ability. *Reading Research Quarterly*, 1969, *4*, 358-365. (a)

Bormuth, J. Empirical determination of the instructional reading level. In Figurel, J. A. (Ed.), *Reading and realism: Proceedings of International Reading Association*, 1969, *13*, part I, 716-721. (b)

Bradley, J. M., Ackerson, G. & Ames, W. S. The reliability of the maze procedure for classroom assessment. *Journal of Reading Behavior*, in press.

Durr, W. R., Windley, V. O. & McCourt, A. A. (Eds.). *Houghton Mifflin Reading Program*. New York: Houghton Mifflin, 1971.

Fry, E. B. A readability formula that saves time. *Journal of Reading*, 1968, *11*, 513-516, 575-578.

Gallant, R. Use of cloze tests as a measure of readability in the primary grades. In J. A. Figurel (Ed.), *Reading and inquiry: Proceedings of International Reading Association*. Newark, Del: International Reading Association, 1965, *10*, 286-287.

Guthrie, J. T. Reading comprehension and syntactic responses in good and poor readers. *Journal of Educational Psychology*, 1973, *65*, 294-300.

Guthrie, J. T., Seifert, M., Burnham, N. A. & Caplan, R. I. The maze technique to assess, monitor reading comprehension. *The Reading Teacher*, 1974, *28*, 161-168.

Jenkinson, M. *Selected processes and difficulties of reading comprehension*. Unpublished doctoral dissertation, University of Chicago, 1957.

Rankin, E. *An evaluation of the cloze procedure as a technique for measuring reading comprehension*. Unpublished doctoral dissertation, University of Michigan, 1957.

Rankin, E. F. & Culhane, J. W. Comparable cloze and multiple-choice comprehension test scores. *Journal of Reading*, 1969, *13*, 193-198.

Ransom, P. E. Determining reading levels of elementary school children by cloze testing. *Proceedings of the International Reading Association*, 1968, *12*, part 1, 477-482.

Taylor, W. L. Recent developments in the use of cloze procedure. *Journalism Quarterly*, 1956, *33*, 42-48.

Taylor, W. L. Cloze readability scores as indices of individual differences in comprehension and aptitude. *Journal of Applied Psychology*, 1957, *41*, 19-26.

Thorndike, R. L. (Ed.). *Educational Measurement*. Washington, D.C.: American Council on Education, 1971.

ELLEN M. WEST
Alachua County Adult Education
Gainesville, Florida

The use of an unobstrusive screening device to approximate reading levels of adults

The need for a minimal threat reading evaluation instrument has been recognized by adult educators for some time (Fay, 1966; Otto & Ford, 1967; Bowren & Zintz, 1977). In recent years, there have been several attempts to devise such an instrument (Smith, Cook, Donaldson, Morrow, Stayman & Bradtmuellar, 1965; White & Kirkland, 1971; Leibert, 1972; Rakes, 1973; Brown, 1974). However, the majority of these attempts have utilized formal testing situations and required the adult student to read orally.

In order to reduce the possibility of interference from test anxiety, an unobtrusive screening device (Webb, Campbell, Swartz, & Sechrest, 1972; Farr, 1970) was developed for use within an adult education center. The device was designed as a student survey and used to approximate reading levels without formal testing.

Scores were expressed in grade levels 0 - 9, with the 9 level indicating a functional adult reader. Specific diagnosis was not the intent of this "testing", but rather a screening process to identify reading levels of entering students.

The device, the *West Informal Reading Evaluation* (WIRE) (West, 1977), was disguised as a student survey and presented in three brief, separate passages and written in relation to adult school and student classes. The passages were written at the 3, 6, and 9 grade levels and arranged in ascending order of difficulty. Readability was determined with the use of the Maginnis (1969) adapted formulation of the Fry Readability Graph (1968). Content of the passages elicited a brief written response (Betts, 1957) from the student, which served as the basis for evaluation.

Method

Subjects

One-hundred fifty-four adults, including both Basic Education students as well as Adult High School Completion students, completed form-1 of the WIRE. One-hundred twenty-three of the same students were then tested using the *Reading for Understanding Test* (RFU) (Thurstone, 1963). The RFU was utilized as the validity criterion and was one of the instruments used in this particular center. One-hundred of the original 154 students were randomly asked to complete a parallel form of the WIRE, the WIRE-2, in order to determine reliability. Students ranged in ages from 16 to 70 with the average age being 25 years. Black and white Americans constituted the largest portion of students with only a small segment of foreign students represented. Approximate distribution of males to females was 40% to 60% respectively.

Procedure

The WIRE-1 was group administered to the students by their classroom teachers. The instrument was distributed as a student survey and all students were encouraged to share their individual ideas, views and opinions.

Distribution, completion and collection of the WIRE took approximately 15 to 20 minutes with a class discussion of the passage topics following the collection of response sheets. The discussion was included in order that all students would be able to express their views whether or not they were able to read all of the passages. Students were told that the survey was also a means of estimating how well they read and that this would enable teachers to select materials, plan classes and help students to register more accurately.

Due to the unobtrusive nature of the instrument, it was anticipated that some adult students would object to being tested without giving their consent. However, no negative reactions were noted and most students were curious as to how well they had scored.

WIRE scores were based upon the correctness, feasibility and completeness of the student's response to each passage. Estimated reading levels were indicated by scores described in the following scale:

0	no apparent reading ability
- 3	below a 3rd grade level reading ability
3	approximately a 3rd grade reading ability
3 +	better than a 3rd grade level, but not more than a high 4th grade level
- 6	below a 6th grade level, but not lower than a high 4th grade level
6	approximately a 6th grade level ability
6 +	better than a 6th grade level, but not more than a high 7th grade level
- 9	lower than a 9th grade level ability, but not lower than a high 7th grade level
9	approximately a 9th grade level ability

Occasionally, readers responded correctly and exceptionally well at the level 6 passage, but entered a "no comment" at the level 9 passage; a score of 6 + seemed to be a low estimate and a -9 possibly a closer approximation of the reader's true ability. Judgment in borderline situations was enhanced with scoring practice and was not as intuitive or ambiguous as it would appear.

Results

Validity coefficients of .868 and .818, respectively, indicated that the *WIRE-1* and *WIRE-2* were both valid measures of adult silent reading comprehension as related to the RFU. Reliability of the *WIRE* scores was determined at the .847 level based on the correlation of 100 paired *WIRE* scores.

In order to determine the inter-rater reliability of the *WIRE*, two teachers (one trained, the other untrained in scoring the *WIRE*) rated 40 randomly selected surveys. Their ratings were correlated with the ratings of the researcher and resulted in coefficients ranging from .823 to .949 which were significant at the .01 level. Table 1 lists the inter-rater means, standard deviations, and range of scores.

The passages of the *WIRE*, a total of 6 for both forms, also elicited important data other than approximate reading levels. Subjective responses provided information that characterized a specific adult education population and lent insight into adult students in general. Areas of inquiry included means of transportation used by the student, educational preferences, vocational goals, and curricular needs as related to career plans.

Discussion

One of the advantages of using an informal method for assessing reading levels was that it could be developed and tailored to fit a specific population.

Table 1

Means, Standard Deviations and Range of Scores of Raters on Forty WIRE-2 Tests

Rater	Mean	Standard Deviation	Range of Scores
1 (Researcher)	7.62	1.46	4.0 - 10.0
2 (Trained)	8.00	1.33	4.0 - 10.0
3 (Untrained)	7.85	1.73	3.0 - 10.0

The *WIRE* was one form of an instrument that could be teacher constructed and adapted to any program or locale. With intervals of 0,3,6, and 9, the *WIRE* was not sensitive enough to distinguish small increments of growth within short periods of time. However, with some adult students who progress rapidly it might be possible to use a form of the *WIRE* every three to four months. The *WIRE* also has potential for assessment of writing, spelling and grammar. However, much more would be needed in the way of effective scoring procedures for accurate analyses.

The use of an unobstrusive screening device, such as the *WIRE*, would be a feasible method of use with adult students, children, and adolescents.

References

Betts, E. A. *Foundation of reading*. New York: American Book Company, 1957.

Bowren, F. R. & Zintz, M. V. *Teaching reading in adult basic education*. Dubuque, Iowa: Wm. C. Brown Company, 1977.

Brown, D. A. *Test that's not a test*. Greeley, Colorado: Basic Education Trade House, 1974.

Farr, R. Use of tests in evaluation: Reading tests and teachers. In D. L. De Boer (Ed.), *Reading diagnosis and evaluation*, Vol. 13, Part 4, Newark, Delaware: International Reading Association, 1970.

Fay, J. B. Psychological characteristics affecting adult learning. In F. Lanning & W. Many (Eds.), *Basic education for the disadvantaged adult*. New York: Houghton Mifflin Company, 1966.

Fry, E. B. The readability formula that saves time. *Journal of Reading*, April 1968, 513-515.

Leibert, R. *Adult informal reading inventory*. Kansas City, Missouri: University of Missouri, 1972.

Maginnis, G. H. The readability graph and informal reading inventories. *The Reading Teacher*, 1969, *22*, 516-518.

Otto, W. & Ford, D. *Teaching adults to read*. Boston, Massachusetts: Houghton Mifflin Company, 1967.

Rakes, T. A. Where to begin. *Adult Leadership*, February 1971, 255.

Smith, E. H., Cook, W. D., Donaldson, W., Morrow, M. J., Stayman, H. R. & Bradtmuellar, W. G. *An informal reading inventory for adults*. Tallahassee, Florida: Florida State Department of Education, 1965.

Thurstone, T. G. *Reading for understanding placement test*. Chicago, Illinois: Science Research Associates, Inc., 1963.

Webb, E. J., Campbell, D. T., Swartz, R. D. & Sechrest, L. *Unobtrusive measures: Non-reactive research in the social sciences*. Chicago, Illinois: Rand-McNally Publishing Company, 1972.

West, E. M. *The use of an unobtrusive screening device to approximate reading levels of adults*. Unpublished doctoral dissertation, University of Florida, 1977.

White, E. & Kirkland, W. *Idaho state penitentiary informal reading inventory*. Boise, Idaho: Boise State College, 1971.

MOLLY M. WILSON
University of Georgia

Future shock and the aged: is reading a cure or part of the problem?

Future shock, Alvin Toffler's term for culture shock in one's own society, is "the shattering stress and disorientation induced in individuals by subjecting them to too much change in too short a time." This rapid rate of change has left all members of society with the problem of adaptability. For the elderly, the shock is in many ways greater than for younger people because they have been witness to many more changes than the younger generations. In addition, time seems to pass more quickly for the elderly and the ability, or perhaps the desire, to adjust to change wanes. For the elderly who are moving from a position as functioning members of society to retirement, adaptation problems may be magnified as they struggle to adapt to both a new life style and a fast moving society. These things, taken together, magnify the effects of future shock on the elderly. Withdrawal, disorientation, and depression are the effects.

While reading is no elixir, it can serve to assuage the adaptation problems of the elderly. Reading can serve to increase consumer awareness among the elderly. Recreational reading can be a method of passing the empty hours of retirement. Paradoxically, however, the very products available for reading contribute in some ways to the adaptation problems of the aged.

Reading as a Cure

Reading Provides Consumer Information

In their adaptation to retirement, the elderly are faced with all types of problems, and the specific ones facing each person depend on the level of sophistication with which he reads. One-fifth of all people over 65 are functional illiterates (Eklund, 1969). This group of aged people suffers the greatest adaptation problems in retirement, for the very business of life requires that one be able to read. The very basis of getting an income in retirement (Social Security, Medicare, and insurance forms) revolves around reading, and being able to read with a fairly high level of proficiency. More courses for the illiterate elderly should be established so that they can function effectively in this highly literate society.

For the average elderly person, reading offers a means for getting consumer information. Haworth (1955) found that the elderly expressed an interest in adult education courses that focused on consumer information. Yet few courses are available that focus directly on the needs and interests of the aged, or that meet during the day when they would feel free to attend.

A small and very select group of the elderly are reading consumers in a very special way. They use the free tuition benefits now available in most states and take college courses. This type of program, which uses reading to provide knowledge for knowlege's sake, appeals to only a small proportion of the elderly and yet it is one of the few services provided uniformly by all states to the elderly. Classes in basic reading instruction for the illiterate and consumer education for the average elderly should receive more emphasis. Haworth (1955) found 70% of the aged expressed educational needs. With such a large expressed need, certainly more can be done.

Reading as a Leisure Activity

Aside from serving an important function as an information service, reading also aids the adaptation of the elderly by serving as a leisure time activity, both a solitary activity and a means for group activity.

The elderly, retired from the working world and suffering some amount of physical disability, have many adjustment problems. Retirees often find themselves leading a secluded life. Inactivity breeds withdrawal, which gerontologists call *disengagement*. Reading is seen as an alternative, a way to keep the mind active and pass hours alone (Cassata, 1967).

Reading groups can serve as a means for staying active. The time spent interacting with others who share an interest in reading can be enjoyable as well as therapeutic. Lovelace (1977) conducted a reading group for residents of a nursing home and found the sessions to be good therapy for participants. Not all the participants were literate, but all viewed the group as a means of remaining active and as something to look forward to. Wilson (1977) conducted a group in a retirement center, comprised of healthy and avid readers. Her findings are similar. The group provided a scheduled weekly activity for the elderly. Reading was actually the springboard for a group social hour, since no reading was done in the meetings, and little reading was discussed. Reading served as the common interest to gather people for social interaction.

So, reading as a leisure activity serves as a useful tool to help overcome adaptation problems. Whether reading alone to pass empty hours, or participating in a group and interacting with others, reading is a possible answer to adaptation problems for the elderly.

Reading as Part of the Problem

Unfortunately the means for a solution serve as part of the problem, also. Written materials are being produced at an astounding rate in this country. Yet the rate of production and the products marketed raise three problems for the elderly.

The Cost of Materials

First, the cost of publishing is skyrocketing, leading to a substantial increase in the price of written materials. For the elderly who have limited incomes, the prices may make owning books prohibitive. Avid readers find a means to get reading material, usually by swapping books with friends or children. But the cost of magazines and newspaper subscriptions makes these materials scarce in the homes of even the most avid readers. New library services such as residential bookmobiles, retirement center libraries, and the free materials for the disabled reader from the Library of Congress are an attempt to make reading materials more readily available. But, are they enough?

Available Materials

Second, one must examine the reading materials that are available and ask if they are of interest to the elderly. The National Survey of Library Services to the Aging (1971) found that most library services focus on children, while less than 2% of all funds are spent on the elderly. The American Library Association publishes only one booklist, *Paths to a Long Life*, that lists books of interest to the elderly. An examination of this list shows no pleasure listings. All books on the list provide consumer information, which is certainly needed, but is not enough in and of itself.

Schramm (1969) found that the elderly spend 45% of their time in media activities. Yet, the median number of hours spent in reading was one hour a day, or about 7% of their total time. So little time in reading — why? Part of the problem may lie in the fact that available materials do not appeal to the elderly, a paradox in a society that is producing written materials faster than ever before in history.

Kingston (1977) also suggests that the nature of materials available might be responsible for the small amount of time the elderly spend in reading. Few magazines appeal to the elderly, e.g., *National Geographic* and *Readers' Digest*. Elderly men cannot identify with the fictional male protagonist who is young and virile. Elderly women prefer light romances with no sex and dislike mystery and science fiction (Romani, 1970). One must ask if these types of materials are readily available. A quick look at the book racks in stores and the fact that U.S. libraries spend so little money on the elderly lead one to suspect the answer is "no."

Vocabulary Change

Third, reading presents another, quite different problem exemplifying some of the problems of future shock. This is a basic comprehension problem brought on by the fast change in vocabulary. Terms like *fastback, wash and wear, flashcube,* and *intercom* appear as new words in reading material. Reading is impeded by new vocabulary, resulting in basic comprehension problems.

While conducting the reading group in Athens, Georgia, this writer was contacted by an elderly woman. An avid reader, she was to read a magazine article and report on it for her mission society. She felt defeated because the article, which was about her church's new evangelistic mission program beaming radio programs via Telstar to Iron Curtain countries, made no sense to her. A quick explanation of the electronic concepts involved in the project easily remedied the comprehension problem, grounded in lack of background experience and vocabulary. This is just one example of what many elderly people who are good readers face when they find themselves bogged down in *Time Magazine*, the newspaper, or even the church missionary magazine.

Where to Go from Here

Adaptation problems exist for the elderly and rise from two sources. First, the shift from being employed and raising a family to retirement causes problems. Second, this change in lifestyle is aggravated by the fast pace and constant change, the future shock, in society. Both factors contribute to the withdrawal and adaptation problems of the elderly.

Reading offers solutions. It can be a means of disseminating information. However, society has fallen short in providing many sorely-needed consumer services or basic education in living for the elderly. Reading also offers a leisure activity. It is a viable means for dealing with both future shock and disengagement in the aged, serving as a solitary activity and as a means for group interaction. Here, again, while inroads are being made through some few agencies, much work has not even begun. The services of the Library of Congress need to be advertised. More reading groups in nursing homes and retirement centers must be established. And certainly a society that produces so many written materials can publish books of interest to the elderly.

Reading is offered as at least a partial solution for the adaptation problems of the elderly. It is time to begin making headway in using reading as a method of helping the elderly adjust to a new lifestyle in a rapidly changing world.

References

Cassata, M. B. A study of the mass communications behavior and the social disengagement behavior of 177 members of the Age Center of New England (Doctoral dissertation, Indiana University, 1967). *Dissertation Abstracts International*, 1968, *28*, 3765A-3766A. (University Microfilms No. 68-2275)

Eklund, L. Aging and the field of education. In M. W. Riley (Ed.), *Aging and society* (Vol. 2). New York: Russell Sage Foundation, 1969.

Haworth, E. H. The expressed educational needs of older-age adults in two senior centers (Doctoral dissertation, Stanford University, 1955). *Dissertation Abstracts International*, 1956, *16*, 60A. (University Microfilms No. 56-96)

Kingston, A. J. Reading and the aged: A statement of the problem. Accepted for publication in *Journal of Educational Gerontology*.

Lovelace, T. Enhancing the lives of nursing home patients through reading activities. Accepted for publication in *Journal of Educational Gerontology*.

National Survey of Library Services to the Aging. Cleveland Public Library. U. S. Department of Health, Education, and Welfare. Office of Education. Bureau of Libraries and Educational Technology, December 1971. (ERIC Document Reproduction Service No. ED 067 521)

Romani, D. Guidelines for library services to the institutionalized aging. *American Libraries*, 1970, *1*, 286-289.

Schramm, W. Aging and mass communication. In M. W. Riley (Ed.), *Aging and society* (Vol. 2). New York: Russell Sage Foundation, 1969.

Toffler, A. *Future shock*. New York: Random House, 1970.

Wilson, M. M. Enhancing the lives of aged in a retirement center through a program of reading. Accepted for publication in *Journal of Educational Gerontology*.

TERRY LOVELACE
University of Southwestern Louisiana

Problems in the measurement of intellectual and linguistic capacities, including reading, in the aged

Experimentation designed to investigate the hypothesis that mental capacity declines with aging has proven troublesome to investigators, many of whom encounter the following pernicious factors:

1) physiological and/or psychological health of subjects;
2) cautiousness;
3) errors in fundamental experimental design.

Psysiological and/or Psychological Health

Is aging a clearly definable series of biological and social events, or is it a process that can be discussed only metaphorically? Not only are there no mechanical, electrical, or chemical processes that parallel aging, but, to complicate the problem, humans have four brains (neuronal, glial, biochemical, and vascular), each with its own genetic history, each vulnerable to specific processes, and each of which may age at a different rate. As a rule, biological characteristics show earlier maturation and decline than psychological capacities and functions. The pattern of age decrement is unique for each person, with rate of decline varying from one function to another, and from one individual to another (Verwoerdt, 1973). There is difficulty distinguishing clearly between biological aging, pathological aging, and disease.

Nevertheless, there are some developmental truisms associated with aging: Increasing age generally is accompanied by a loss in sensory acuity, resulting in information being received incorrectly or partially, without the subject being aware of it. Perceptions of changes in body position and orientation in space become less accurate with age. A loss of strength and agility is noted. Reaction time begins slowing shortly after age 20; movement time and spontaneous speech, however, do not decline until after age 70. Unfamiliarity and/or extraneous stimulation slow down the learning of new tasks. Pressure for speed results in the slowing of a task and in higher error rate (Carp, 1973).

However, before these developmental truisms are applied to *all* elderly persons, researchers should consider the major conclusion reached by Goldstein and Shelly (1975), who examined the similarities and differences between psychological deficit in aging and brain damage. They suggest that one cannot speak of mental decline with aging in a global way, and, that if trends toward declines with age do exist, the conclusion still cannot be drawn that a decline with age exists in all intellectual functions.

The future improvement of measurement of intellectual functioning among the aged will be dependent to a large extent upon investigators' abilities to separate contextual effects from essentially age-related changes, i.e., physiological and psychological changes. One possibility for more closely defining age of persons is the Halstead battery of neuropsychological tests–held to be reliabale and valid in defining the state of higher-brain functioning independently of calendar age, sex, formal education, ethnic origin, occupational history, and psychometric intelligence (I.Q.). The nonverbal tests attempt to determine the brain's ability to integrate in time and space information supplied to it by vision, hearing, and touch (Halstead, 1963).

The loss of perceptual acuteness may have led researchers to erroneous generalizations as to the loss of cognitive abilities among the aged. Meyerson (1976) suggests that, with the exception of the common finding of prebycusis (deterioration in hearing), the processes of communication do not seem to be seriously affected by normal aging changes. This theory is supported by Granick, Kleban, and Weiss (1976), who investigated the relationships between mild hearing loss and cognitive functioning for two independently selected samples of aged persons whose hearing was within normal limits. The findings imply that aged persons may be more intellectually capable than their test performances suggest, and that hearing is an important variable to be considered in the assessment of their cognitive functioning.

The possibility that reduced effectiveness in cognitive functioning is associated with reductions in auditory acuity in persons who can hear and respond appropriately to normal speech raises significant questions about the reliability, validity, and meaning of much of the currently-available gerontological literature. It is unusual, for example, for the hearing status of aged persons to be reported in descriptions of learning experiments, despite the fact that the stimulus materials are often orally resented. The same holds for other types of investigations which deal with intellectual factors, reaction time, and personality functioning wherein auditory perception and discrimination are involved.

Maddox and Douglas (1974) suggest that dispersion of scores on a variety of physiological indicators increases over time for a cohort of older individuals. One of the most severe aging effects is labeled *terminal drop.* Terminal drop (rapid change) in social, psychological, and physiological functioning immediately prior to death has been suggested by several writers as a possible source of individual differences in late life (see Riegel & Riegel, 1972, for further citations). In studying elderly persons, it is argued, investigators are in fact studying samples which include both elite survivors and an undertermined proportion of individuals experiencing rapid decline in functional capacity prior to death. Carp (1973) supports this argument in regard to I.Q. scores, which are known to drop with serious illness. In some sense, the decline in I.Q. with age is an artifact; if those people who are within a year of death are eliminated from the average, intelligence test scores do not change significantly.

Problems in measuring cognitive and linguistic functioning are often compounded by the gradual onset of impaired orientation, judgment, and intellectual functioning seen in the elderly due to *senile dementia.* It has been estimated that as many as 10% of the elderly demonstrate symptoms of senility, the frequency being 20% above the age of 80 (Aire, 1973). Unfortunately, as Haase (1971) remarks, several diseases associated with the elderly are clinically indistinguishable from *senile dementia.* If researchers do indeed attempt to match persons on the basis of health status, great care is needed to avoid crediting the elderly with loss of function due to senility when senility is actually not present.

Cautiousness

Birren (1974) argues that with age the central nervous system slows its capacity to take in, store, and retrieve information, which presumably influences a wide range of behaviors. But many researchers refute this theory, and align themselves with Botwinick (1969), who concluded that what appears to be a decline in retrieval speed may be instead cautiousness on the part of the elderly. Cautiousness in this case being defined as a hesitancy to become involved when risk is high. An experiment involving two semantic memory tasks designed by Eysenck (1975) revealed a differential effect of age on recall and recognition, the older persons (12 working schoolteachers, CA = 55-65) responding more slowly on the recognition task, but not on the recall task. The response times of the older persons were less affected by the dominance of the to-be-retrieved information than were the younger persons (12 students at the University of London, CA = 18-30). These results suggested that persons in the older age group may have retrieved information faster than the young persons, but that they required longer to decide upon a response, i.e., they were more cautious. A replication of Eysneck's experiment with a larger sample is suggested, as generalizations from his study are necessarily limited.

In an experiment involving older (CA = 69.31) and younger (CA = 21.71) males participating in a vocabulary task involving various degrees of risk, older males were found more cautious than young adults. The older males were also less likely to raise their aspirations following success. No effect of instructions on cautiousness was found for either age group (Okun & Divesta, 1976). In a similar study, Birkhill and Schaie (1975) investigated the effects of differential reinforcement of cautiousness on intellectual performance in the elderly, employing 56 females and 32 males with a mean age of 73 years. Cautiousness was manipulated by systematically varying pre-test instruction conditions involving the reinforcement of two levels of risk and of response omission when taking the Primary Mental Abilities Test. Results suggest that performance in intelligence tests is much influenced by situational variables involving motivational components. Persons exposed to low-risk conditions performed significantly better on the cognitive tasks than those exposed to high-risk conditions only when they were able to exercise the option of responding or not responding on individual test items.

Problems with Fundamental Experimental Design

Experimental behavioral studies of aging are either cross-sectional (stratified) or longitudinal. Each design has its own strengths and weaknesses. In the longitudinal study the individual is his own control. The cross-sectional approach is particularly vulnerable to two factors: 1) the cryptic, or hidden, variable, and 2) sampling biases followed by overgeneralization. An example of the cryptic variable is cited in the study by Jones and Conrad (1933), who applied a modification of the Army Alpha examination from World War I to all adults between the ages of twenty and sixty in a New England community. The decrement in psychometric intelligence after thirty was attributed by the authors to aging, but Lorge (1956) has cast doubt on the validity of this interpretation by showing a probable inverse relation between the person's formal schooling and calendar age.

The cross-sectional approach is particularly susceptible to the problems of purposive non-random sampling, wherein persons are selected by some purposive method, i.e., availability, and are therefore subject to the biases of personal selection. All too often, stratified cross-sectional random sampling becomes stratified purposive sampling, e.g., 30 available college freshmen are compared with 30 available elderly in a nursing home. Marascuilo (1971) rather forcefully emphasizes that 1) currently employed statistical methods are dependent upon random sampling and the assumptions underlying this type of sampling, and 2) if the assumptions of this sampling procedure are not met, many statistical analyses cannot be used.

Abrahams, Hoyer, Elias, and Bradigan (1957) summarized psychological research on human subjects published in the *Journal of Gerontology* from

209

1963 through 1974 along a variety of dimensions, including subject selection procedures, sample characteristics, data analysis techniques, research design, author characteristics, and area of psychological research. Cross-sectional design strategies were found in 67.2% of the studies published in the 11-year period, which augurs ill for the generalizations made from these studies. Thirty-seven per cent of studies published during the 1969-1974 period did not mention the health of the subjects, while only 16.4% controlled for health.

Conclusions

Apparent loss among the aged of intellectual and linguistic capacities, including reading ability, may be due to loss of perceptual input, disease processes, acquired cautiousness, psychological and environmental problems, and/or faulty research design followed by overgeneralization. Suggestions for subsequent experimentation which are designed to counterbalance these deficiencies include:

1) obtaining, if at all possible, a stratified *random* sample for cross-sectional studies;

2) specifying clearly the age characteristics of samples under investigation;

3) undertaking some form of health assessment where it is known or suspected that disease states influence age-associated performance;

4) treating gender differences as sources of variance;

5) where appropriate, making use of multivariate procedures as protection against chance differences resulting from application of univariate ANOVA's to multiple dependent factors;

6) providing an estimate of the proportion of variance accounted for in the dependent variable by the independent variable;

7) generalizing only to the specific population from which subjects are drawn.

Intellectual/cognitive functioning, including reading ability, has continued as the dominant area of research in gerontological psychology. This no doubt reflects the large number of abilities and skills included by the term intellectual/cognitive functioning, the large number of measurement instruments available, and the large number of variables known to affect intellectual and cognitive performance. As the data base and the level of methodological sophistication have increased, the parameters of adult intellectual performance, memory, and problem-solving have become more clearly delineated; but in view of increased knowledge, there has also been a need to reexamine previous assumptions based on limited knowledge and faulty design.

References

Abrahams, J. P., Hoyer, W. J., Elias, M. F. & Bradigan, B. Gerontological research in psychology. *Journal of Gerontology*, 1963-1974: Perspectives and progress. *Journal of Gerontology*, 1975, *30*, 668-673.

Aire, T. Dementia in the elderly: Diagnosis and assessment. *British Medical Journal*, 1973, *4*, 540-543.

Birren, J. F. Translations in gerontology–from lab to life: Psychophysiology and speed of response. *American Psychologist*, 1974, *29*, 808-815.

Birkhill, W. R. & Schaie, K. W. The effect of differential reinforcement of cautiousness in intellectual performance among the elderly. *Journal of Gerontology*, 1975, *30*, 578-583.

Botwinick, J. Disinclination to venture response versus cautiousness in responding: Age differences. *Journal of Genetic Psychology*, 1969, *115*, 55-62.

Carp, F. M. The psychology of aging. In R. R. Boyd & C. G. Oakes (Eds.), *Foundations of practical gerontology*. Columbia, S. C: USC Press, 1973.

Eysenck, M. W. Retrieval from semantic memory as a function of age. *Journal of Gerontology*, 1975, *30*, 174-180.

210

Goldstein, G. & Shelly, C. H. Similarities and differences between psychological deficit in aging and brain damage. *Journal of Gerontology*, 1975, *30*, 448-455.

Granick, S., Klaban, M. H. & Weiss, A. D. Relationship between hearing loss and cognition in normally hearing aged persons. *Journal of Gerontology*, 1976, *31*, 434-440.

Haase, G. R. Diseases presented as dementia. In C. E. Wells (Ed.), *Dementia*. Philadelphia, PA: F. A. Davis, 1971.

Halstead, W. C., Merryman, P. & Klien, B. The Halstead index and differential aging. In R. H. Williams, C. Tibbits & W. Donahue (Eds.), *Processing of aging: Social and psychological perspective* (Vol. 1). New York: Atherton Press, 1963.

Maddox, G. L. & Douglass, E. B. Aging and individual differences: A longitudinal analysis of social, psychological, and physiological indicators. *Journal of Gerontology*, 1974, *29*, 55-563.

Marascuilo, L. A. *Statistical methods for behavioral science research*. New York: McGraw-Hill, 1971.

Meyerson, M. D. The effects of aging on communication. *Journal of Gerontology*, 1976, *31*, 29-38.

Okun, M. A. & DiVesta, F. J. Cautiousness in adulthood as a function of age and instructions. *Journal of Gerontology*, 1976, *31*, 571-576.

Riegel, K. F. & Fiegel, R. M. Development, drop and death. *Developmental Psychology*, 1972, *6*, 306-319.

Verwoerdt, A. Biological characteristics of the elderly. In R. R. Boyd & C. Oakes (Eds.), *Foundations of practical gerontology* (Rev. ed.). Columbia, S. C.: USC Press, 1973.

PATRICIA L. ANDERS
DOUGLAS CARDELL
University of Arizona

Self-perceptions and reading habits of adolescents, college students, and adults

A popular conception exists that people could read if they chose to do so, but that, for the most part, they do not choose to do so. Mikulecky reported three surveys in the 1977 *NRC Yearbook* which bear witness to this idea: one survey indicated that 1% of the adult population completes a book per year; a second survey suggested that 25% of all adults read in a book per month; and a third survey reported that 58% of the population "never read, never finished" a book.

This assertion and these surveys provided the impetus for this study. Three randomly selected groups were chosen: (1) adult readers; (2) college students; and (3) junior high students. The subjects were surveyed to examine and compare their reading behaviors within three parameters: (1) their self-perceived reading abilities; (2) their self-reported reading habits; and (3) their ability to answer an objective-type measure of reading retention. By examining and comparing these three groups, the assertion that the general population does not choose to read could be rejected or supported, at least for people across three age groups in Tucson, Arizona.

Method

Subjects

The adult population sample was chosen by randomly selecting 209 names from the local telephone directory. When telephoned, 199 of these 209 potential subjects agreed to respond to a reading survey. Demographic data reported by the 199 respondents matched, within 1%, published census data with regard to age, income, education, and ethnicity.

The college students were surveyed with questionnaires distributed in randomly selected liberal arts classes. Questionnaires were completed by 176 subjects. The average age of this group was 20 years, the average family income was $15,500 and the average education was 14 years.

The junior high students were also sampled with questionnaires handed out in school classes. Two classes were randomly selected from each of four public Tucson junior high schools. A total of 199 students completed the questionnaire. This sample also closely corresponded with published demographic data.

Materials

All subjects were asked the same questions whether on the telephone or on a dittoed handout. The questions were organized into four different types. The first type related to reading habits: how often newspapers, news magazines, trade magazines, and fiction and non-fiction books were read, and frequency of trips to the library.

The second type of questions related to the subjects' self-perceptions of their reading abilities as compared with others about their own age. The subjects responded to four statements by choosing one of three degree-related statements (i.e., somewhat more, about the same, somewhat less, etc.).

The third type of questions were designed to ascertain demographic characteristics of the subjects: age, income, education, sex, and race were included as demographic data.

The final category of questions was a true-false quiz to test retention of newspaper information. Questions were chosen for their suitability from the headlines of the two Tucson newspapers from the previous week. One or two words from the headline were then changed, added, or omitted in such a manner as to give the new headline an opposite or greatly altered meaning. For example: "City council votes to raise water rates" became "City council votes to lower water rates."

Analysis

The analysis consisted of (1) Pearson-Product-Moment Correlation coefficients to compare variables across all three groups, (2) calculation of each group's mean responses to each of the items, and (3) *t* tests for differences (a two-tailed test) between each of the means.

Results

The amount of reading reported by each of these three groups was a good deal greater than predicted (see Table 1). Library usage also appeared to be a fairly common habit across all three groups: college students reported going to the library an average of ten times per month, junior high students reported an average of seven times per month, and the adult population reported an average of three times per month. Both the college students and the junior high students made significantly more trips to the library than did the adult population, $t(370) = 12.64, p < .01$, and $t(394) = 5.29, p < .01$, respectively.

Table 1

Mean Number of Reading Materials Read Per Year

Group	Newspapers	News Magazines	Trade Magazines	Fiction	Non-Fiction
General Population	312	32.4	28.8	21	15
College Students	395.2	39.6	9.6	15	19
Junior-High Students	254.8	25.2	6.0	59	35

No significant differences were found between the adult sample and the college sample with regard to newspaper readership. However, the adult population reported reading the newspaper significantly more than did the junior high students, $t(396) = 4.19, p < .01$.

The quantity of newspaper, fiction, and non-fiction reading correlated with each of the group's attitudes toward reading (see Table 2).

Table 2

Pearson Product Moment Correlations Between Attitudes and Type of Reading

Attitude		*Type of Reading*		
	n	Newspaper	Fiction	Non-Fiction
Rate				
Adults	199	–.2388***	–.2405***	–.1726**
College	176	–.2171*	–.1556*	–.1523*
Junior High	199	–.2966***	–.3223***	–.2835***
Comprehension				
Adults	199	–.1230*	–.2474***	–.1627**
College	176	–.1298*	–.2515***	–.1684**
Junior High	199	–.2349***	–.0752	–.1065
Enjoyment				
Adults	199	–.2405***	–.3143***	–.1647**
College	176	–.1269*	–.1702**	–.0325
Junior High	199	–.2713***	–.2475***	–.2784***
Amount				
Adults	199	–.2288***	–.3152***	–.1543*
College	176	–.1682**	–.2262***	–.1537*
Junior High	199	–.1382*	–.3277***	–.3323***

*$p < .05$
**$p < .01$
***$p < .001$

Furthermore, college students thought they read more than other people their own age significantly more, $t(369) = 2.53$, $p < .024$, than did the adults. Junior high students reported reading faster than people their own age significantly more than did the adults, $t(396) = 3.75$, $p < .01$.

Finally, no significant differences were found between the population groups in their scores on the quiz designed to estimate reading retention. All groups averaged about 58% correct responses.

Conclusions

Tucsonans reported reading far more than the previously reported studies would suggest. This result indicates that perhaps people are reading more than has been previously thought.

Also, across all age groups, the subjects' self perceptions about reading and their reported reading behaviors were very similar. This tends to support general folkloric notions that "either people do well at what they enjoy or they enjoy what they do well," even to the point of intellectual behaviors such as reading.

References

Arizona Statistical Review (32nd ed.). Valley National Bank, Tucson, Arizona, 1976.

Mikulecky, L. J. & Wolf, A. Effect of uninterrupted sustained silent reading and of reading games on changes in secondary student's reading attitudes. In P. D. Pearson (Ed.), *Reading: Theory, research and practice* (Twenty-sixth yearbook of the National Reading Conference). Clemson, South Carolina: NRC, 1977.

GEORGE W. McCONKIE
Cornell University

Stimulus control in the study of reading

The act of reading is a very complex process, as many researchers have noted, and can be studied at a number of different levels. The purpose of this paper is to describe a research approach being taken at Cornell University to study reading at the micro-level. This project is aimed at gaining an understanding of the perceptual and cognitive events occurring during a fixation and across successive fixations during reading by relatively skilled readers. This paper will describe the technology which has been developed for the research, pose a number of questions about what is taking place during reading, give examples of studies bearing on these questions, and discuss the potential value of the research approach being taken.

Technology

Over the past several years an eye-movement contingent display control system has been developed for the study of reading. This is a computerbased system which monitors a reader's eye movements as he is reading from text displayed on a cathode-ray tube (CRT), and which is capable of making changes in the text contingent upon characteristics of his eye movements. The details of this system have been described elsewhere (McConkie, Zola, Wolverton & Burns, in press), and will only be briefly mentioned here. The system checks the location of the eye 1000 times per second, and keeps a complete record of eye movements during reading. In addition, it detects when a saccadic eye movement begins and ends, and calculates where the eye is directed on the CRT (on what letter position the eye is centered). As the computer obtains this eye movement information during reading, it is able to make changes in the text while reading is in progress, based on whether the eye is moving or still, or where the eye is directed. For instance, a change in the text may be produced during a single eye movement so the text is different in some specific way on one fixation than it had been on the last. Changes can also be made during a fixation, so that a certain number of msec. after a fixation begins one line of text is removed and quickly replaced by another. The nature of the stimulus change can depend on where the eye is centered. At the present time we are developing a prediction algorithm by which the computer can calculate during an eye movement where the eye is going to stop, based on the velocity pattern of the eye, and make changes in the display so that when the eye comes to rest at that location certain stimulus patterns will lie at specific foveal or peripheral visual locations for that fixation. These display changes can be produced very rapidly: replacing one 70-character line with another requires 2 msec. The nature and use of this computer system will become clearer as we proceed through some specific research examples. This will be done by bringing up a series of questions about reading, and dealing with each in turn.

Some Examples of Research

What information is carried from one fixation to the next?

The first study to be described was carried out by Rayner (1975) several years ago using a system which we developed at the Massachusetts Institute of Technology (McConkie & Rayner, 1974). In the study subjects read short paragraphs, each containing a specific word location of interest, called the critical word location. As the subjects read, an eye movement was detected which was likely to terminate in a fixation centered on that word location. The contents of that location (a word or a non-word letter string) were immediately replaced by a word of the same length. Thus, the word which occupied that location when the eye landed there was different than the word that had occupied it on the prior fixation. If the text prior to the change read "The robbers guarded the palace. . .", on the next fixation the text read "The robbers guarded the police. . .". The word "palace" had been replaced by the word "police" while the eye was moving to look at that word. Two questions of interest are: Would a reader detect such a change, and if so, when would it be detected? Several theories of reading suggest that visual information is acquired and stored in some sort of buffer storage during a fixation, and only used at a later time when needed for reading (Shebilske, 1975). If this were the case, it might be expected that the discrepancy would not be discovered on the fixation following the change (when the new word was being seen for the first time), but only during a later fixation. It is of interest to find, then, that there was a substantial increase in the duration of the fixation immediately following the change. Thus, the discrepancy in the text pattern from one fixation to the next was detected in some manner, and this occurred during the fixation in which the new stimulus was first seen. It is not true that there is a considerable lag between the input of visual information and the time it can be involved in the reading process. It is also apparent that the information acquired during a fixation can influence the duration of that fixation. Finally, it is of interest that when the subject's fixation prior to the change was within three letter positions in front of the critical word location (thus close enough that the word could probably be read), there was no evidence on the next fixation that the change had been detected even though the eye was centered on or immediately following the changed word. Thus, although the word was clearly within the visual region on which it could be read if desired, it was being ignored. Apparently not all words within the visual span region are "seen" during a fixation.

The Rayner study just described clearly shows that information is carried over from one fixation to the next; otherwise the change in the text between the two fixations would not be detected. McConkie & Zola (in preparation) investigated what information is being carried across fixations. Rayner (1975) suggested that there may be some integrative visual buffer, where the visual information obtained on successive fixations is integrated into a single set of visual information about the text. Display changes from one fixation to the next would disrupt this integration process because visual patterns in certain regions would not match. If this were the case, what would happen if people read text printed in AlTeRnAtInG cAsE, and then during some saccade the case of every letter were changed; that is, every upper case letter were changed to lower case, and vice versa. If the visual pattern from successive fixations is indeed integrated in some buffer, such a change should be highly disruptive to reading. With every letter changing shape from one fixation to the next any attempt at integration of the two visual patterns should be devastating. However, McConkie & Zola (in preparation) found that such changes had no effect on reading. In fact, when subjects knew that such changes were taking place, and tried to see them as they read, they were unable to. Whatever information is carried over from one fixation to the next, it is apparently not simple visual patterns or even shapes of letters or words. It must be information at a deeper level of encoding: perhaps letters, letter pattern, sound patterns, etc. It is also probably not only lexical meanings, since Rayner (1975) provided evidence that display changes are detected at points where subjects are not distinguishing between words and non-words. The nature of the information carried over from one fixation to the next by a skilled reader remains an interesting problem for further investigation.

From what region is visual information acquired during a fixation?

Turning now to another question about reading, a series of studies has been conducted investigating the size of the perceptual span, or the region from which visual information is acquired during a fixation in reading. In a recently-reported study (McConkie & Rayner, 1976), the text display was changed on each fixation. When the computer detected that a fixation had begun, it replaced all letters which lay more than a certain number of character positions to the left and right of the fixation point with other letters. Thus, on each fixation, a stimulus pattern was created such that normal text lay in a specific region around the fixation point, and non-readable letter strings lay outside this region. The following lines show how a line of text might have looked on each fixation as a person was reading part of the line:

I. xxxxxxxxxxx. It appears that pitch-naxxxx xxxxxxx xxx xx
II. xxxxxxxxxxx. Xx xxxxxrs that pitch-naming abilxxx xxx xx
III. xxxxxxxxxxx. Xx xxxxxxx xxat pitch-naming ability cxx xx

On each fixation, the line of text was changed so that the normal letters were present for 4 letter positions to the left of the fixation point and a considerable distance to the right. All letters outside this region were replaced by other letters. The question asked in the study was this: what would happen if the reader could not obtain useful visual information from words or letters more than 4 letter positions to the left of the fixation point? The answer is: nothing happens. Mutilating the text in this region seems to have no effect on reading. However, if letters more than 4 letter positions to the right of the fixation point are replaced by other letters, reading deteriorates dramatically. Apparently the region from which visual information is acquired during a fixation by skilled readers of English is highly asymmetrical, lying primarily to the right of the center of vision. In line with a comment made earlier, this indicates that visual information well within the retinal region from which letters and words can be identified (Bouma, 1973) may be ignored during a fixation in reading. Other information about the size of the perceptual span is reported by McConkie & Rayner (1975) and by Rayner (1975).

How is the eye guided?

Another question of some interest is whether the eye is sent to a specific, preselected location for the next fixation, or whether it is simply advanced a certain distance rather automatically (Rayner & McConkie, 1976; Rayner, 1978, Levy-Schoen & O'Regan, 1977). McConkie & Zola (in preparation) investigated this question by changing the display during a saccade to cause the next fixation to lie at a location different from that to which it was originally directed. This was done by simply shifting the text two letter positions to the right or the left during certain saccades. Thus, on the following fixations the eye was centered on a location two letter positions away from where it would normally have landed in the text. Two facts are apparent from this study. First, subjects are not consciously aware that any change has been made in the display. But second, their eye movement patterns are affected by the shift in the text. There is a definite increase in the number of short eye movements, around two letter positions in length, following text shifts, thus apparently indicating some sort of correction for the erroneous eye position. This suggests that in some way the mind is directing the eye rather specifically, and that the eye movement system detects inaccurate positioning of the eye in reading.

Summary

While the study of reading using eye-movement contingent display control has only begun, the power of this approach is already becoming evident through studies like those described above. In the few studies mentioned here, the following facts about reading seem to be emerging:

1. While visual information about the shapes of letters and words must be acquired during each fixation and used to support reading, such specific visual detail is not carried across from one fixation to the next. This integration across fixations, by which peripheral visual information from one fixation can facilitate processing on the next (Rayner, McConkie & Ehrlich, in press) must be occurring at some deeper level of encoding.
2. During a fixation in reading visual information is apparently derived from specific text locations (or retinal regions) rather than from all of the text lying within the normal perceptual span. Some words lying within the visual field at locations where word identification is possible are ignored during a fixation. This is particularly likely for words lying to the left of the center of vision, from which meaning has probably already been derived during a prior fixation.
3. When a saccade is initiated, the eye is being sent to a specific location in the text. Landing at a position different from this can trigger corrective eye movements of some sort.

The Potential of Research Involving Eye Movement Contingent Control Over the Display

In another paper (McConkie, 1977) an attempt has been made to integrate these and other facts about reading into a way of conceptualizing the relationship between eye movements and language processing in reading. The point to be made here is that the problem of how to exercise precise control over the stimulus as a person is engaged in reading has been largely solved (although refinements are certainly needed) and that such control opens the door to a powerful new approach to the study of reading. Many specific questions about the perceptual and language processing occurring during reading which were previously beyond our ability to study in a precise way become tractable within this research context. There have been many puzzling questions about reading which we have been unable to investigate in a carefully controlled manner. For instance, just what aspects of the text lying within the foveal or near-foveal region are actually being detected and used during a fixation in reading, and how is this influenced by the linguistic context or the reader's expectations at that point in the text? When a new fixation begins, how does the reader know where, in the large visual array, he should seek information: from which line and which region on that line? How long does it take from the time a given region of text (say, a particular word) is visually encountered during a fixation until the language processes have detected or produced certain meaning relationships which involve information from this word? Once alternative answers to such questions as these have been stated, they can often be rephrased into sets of contrasting predictions about what would happen if this or that display or text manipulation were made while a person were reading. Typically these predictions indicate which types of visual or textual manipulations would be detected or would cause difficulty in reading and which would not, or when such difficulties would be encountered, if the perceptual and language processing occurred in a certain way. Research already carried out has shown that the eye movement patterns can be sensitive indicators of the effects of such manipulations on reading (Rayner, 1975; McConkie & Zola, in preparation). Processing difficulties can be detected as changes in the duration of a particular fixation or length of a particular saccade, when averaged over subjects or over trials for the same subject. This is a type of precision which we have needed but largely lacked in our prior reading research.

The studies which have been described here all relate to the nature of relatively skilled reading. It should be obvious that as knowledge of skilled reading develops, these same techniques can be used to study the development of reading skill. But of even greater importance is the fact that, as we come to understand skilled reading better, a firmer basis will be laid for exploring in detail the problems being encountered by less able readers. At some point, the technology which permits such flexible control over the stimulus during reading will become the basis for a new generation of computerized diagnostic instruments which are capable of precisely specifying a large number of very particular aspects of a given person's visual and language processing characteristics from data collected as he is reading.

Certainly there are other levels at which the reading process must be studied, such as those which will indicate what people retain from their reading, and how this is influenced by discourse characteristics, by the nature of the task they engaged in, and by their own prior knowledge. But for studying the reading processes at the micro level which has been described in this paper, the eye movement contingent display control techniques provide a significant breakthrough in the development of reading research capability.

---

1. The preparation of this paper has been supported in part by Grant No. MH24241 from the National Institute of Mental Health. It was presented at the National Reading Conference, New Orleans, Louisiana, in December, 1977. The author wishes to acknowledge the assistance of David Zola, Gary S. Wolverton and David D. Burns in developing the computer system described here, and in conducting much of the research. Requests for reprints should be addressed to George W. McConkie, Department of Education, Cornell University, Ithaca, New York, USA, 14853

References

Bouma, H. Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research*, 1973, *13*, 273-284.

Geyer, J. J. Models of perceptual processes in reading. In H. Singer & R. B. Ruddell (Eds.), *Theoretical Models and Processes of Reading*. Newark, Delaware: International Reading Association, 1970.

Gough, P. B. One second or reading. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by Ear and by Eye*. Cambridge, Mass.: MIT Press, 1972.

Levy-Schoen, A. & O'Regan, K. *The control of eye movements in reading*. Paper presented at the Conference on the Proçessing of Visible Language, Eindhoven, The Netherlands, September 1977.

Matin, E. Saccadic suppression. *Psychological Bulletin*, 1974, *81*, 899-917.

McConkie, G. W. *On the role and control of eye movements in reading*. Paper presented at the Conference on Processing of Visible Language, Eindhoven, The Netherlands, September 1977.

McConkie, G. W. & Rayner, K. *Identifying the span of the effective stimulus in reading*. Final report, U. S. Office of Education Grant No. OEG-2-71-0531, July 1974.

MCConkie, G. W. & Rayner, K. The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, 1975, *17*, 578-586.

McConkie, G. W. & Rayner, K. Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society*, 1976, *8*, 365-368.

McConkie, G. W. & Zola, D. *Is visual information integrated across successive fixations in reading?* Unpublished manuscript.

McConkie, G. W. & Zola, D. *How precise is eye guidance in reading?* Unpublished manuscript.

McConkie, G. W., Zola, D., Wolverton, G. S. & Burns, D. D. Eye movement contingent display control in studying reading. *Behavior Research Methods & Instrumentation*, in press.

Rayner, K. The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 1975, *7*, 65-81

Rayner, K. Foveal and parafoveal cues in reading. In J. Requin (Ed.), *Attention and Performance VII*. New York: Academic Press, 1978.

Rayner, K. & McConkie, G. W. What guides a reader's eye movements? *Vision Research*, 1976, *16*, 829-837.

Rayner, K., McConkie, G. W. & Ehrlich, S. Eye movements and integrating information over fixations. *Journal of Experimental Psychology: Human Perception and Performance*, in press.

Shebilske, W. Reading eye movements from an information-processing point of view. In D. Massaro (Ed.), *Understanding Language*. New York: Academic Press, 1975.

R. P. McDERMOTT
The Rockefeller University

Some reasons for focusing on classrooms in reading research

During the past decade, reading research has mushroomed in volume and sophistication. Much of this research has been centered on what a person must know in order to acquire information from a printed page. This has led to a number of developments, some of them startling, in our sense of what readers must know to process letters (Estes, 1975), words (Smith and Kleiman, in press), sentences (Wanat, 1976), and even texts (Fredericksen, in press; Griffin, 1977). A distant hope underlying all this research is that we might be able to locate the gaps in what deficient readers know about how to read, and that we might use all this research to develop a pedagogy designed to fill these gaps.

As exciting as this progress has been, and as much as we must hope that it can continue, this paper represents a call for an alternative approach to the study of reading. Rather than focus on skills children must develop in achieving reading competence, this paper calls for an inquiry into one of the crucial environments in which children are asked to develop skills, namely, classrooms. The effects of other environments on learning to read, for example, homes, streets and work places, are equally important (Sticht, 1975; Szwed, 1977), but will not be addressed in this paper. As a spur to an inquiry into reading in classrooms, the paper claims that given some uniform organizational constraints on most teachers and children in American classrooms, most particularly, pressures to sort children into categories such as successful and unsuccessful, able and disabled, smart and dumb, different children get differential access to reading instruction and differential time on reading tasks. The product of these organizational constraints is that many children, although perfectly well equipped to learn how to read under other circumstances, fail to develop the reading skills which mark a successful school career. If this conclusion is correct, then our problem with deficient readers is not that they *cannot* develop various reading skills, but that they are not offered appropriate institutional circumstances for developing such skills.

This is a point of considerable consequence for our current strategies of research and pedagogy. At the present time, it is only possible to say how we cannot, in principle, solve reading problems with the findings of present research efforts. It would be nice to offer a more positive statement. Although there are strong signs that reading deficiencies are socially organized, research on how this state of affairs is accomplished with each new generation of school children is just beginning. A recent review of the literature (Cazden, in press) correctly complains that we know little about the conditions under which our children are asked to learn to read. We will have to do better.

If our goals are to understand how some people learn to read while others do not and then to develop an appropriate pedagogy for those who do not, most current research strategies which emphasize the analysis of information processing skills are inadequate. Such research may tell us about what most readers might do while reading but little about how they develop such habits or what would have to be changed in order for them to do so. Such research reveals little about the lives of readers and the organization of the world which keeps them engaged with print.

The first limit on current research into reading is methodological. At the present time, there is no methodological warrant for generalizing from laboratory research on what readers seem to know how to do to the behavior of persons outside the laboratory. Happenings behind the eyes are hard to see. Accordingly, analysts have had to set up well-specified and carefully controlled environments in order to allow for carefully drawn inferences about the information processing skills of their subjects. In being so careful, experimentalists have effectively restricted their generalizing powers to conclusions about what it is their subjects know *only* under the conditions that the experimenter (most likely) controls (Scribner, 1975). New conditions demand new experiments, *ad infinitum*. As such, there are no principled ways of generalizing to the behavior of subjects in everyday life in which the environments for their behavior are not so easily discovered by the analyst. This fact leads to continual surprises for the laboratory oriented investigator, as students of the social and cultural contexts for behavior keep producing accounts of the apparent use of skills in everyday life on the parts of people who do not demonstrate such skills on tests or experiments (Cole, Gay, Glick, & Sharp, 1971; Cole, Sharp & Lave, 1976).

Although we have available many complaints that laboratory research is often ecologically invalid and difficult to use as a basis for policy decisions (Bronfenbrenner, 1977), it has only recently become clear that such research can be systematically and inherently irrelevant to the everyday lives of its subjects. The analytic power allowed by experimental procedures by virtue of their careful control over only certain predefined dimensions of how people organize their behavior renders experimental results nongeneralizable in the face of everyday life scenes which are organized in more complex ways (Cole, Hood & McDermott, 1978).

The second limit of research on skills readers or potential readers must have in their heads is theoretical. It is not clear what role, if any, the development of such skills has in the development of reading competence. It is not even clear that such skills are hard to develop, and it is not fair to assume that the problem with problem readers is that they are missing skills in their heads. Although it is usually assumed that non-readers or poor readers are missing skills which we must teach them with a bag-of-tricks pedagogy, it has not been shown that the presence or absence of such skills is the key to any person's reading problem.

If research is to produce relevant implications for instruction, it must speak to the experiences of the people who have learned how to read or who are learning how to read. Research into the cognitive skills underlying efficient reading is assumed to capture something important about learning to read, but it does not speak at all to the circumstances under which people either learn to read or remain illiterate despite years of instruction. Our current preoccupation with individual skills offers us a theoretical naivete in the face of accounts of how displays of literacy skills appear to be arranged easily under certain types of social organization.

For example, Soderbergh (1973) reports that she can consistently train two-year-olds to perform some basic reading tasks. Conklin (1949) reports a stunning case of an indigenous literacy system from the Philippines in which adolescents acquire literacy skills in the course of a few months. Reports of this kind are not unusual, the most recent being that of Scribner and Cole (1978a,b) who report on a delightful case from West Africa of a three-part literacy (English, Arabic and Vai, with three different orthographies) for many of the members of the Vai tribe.

These examples can be used to show that the acquisition of reading competence, like the acquisition of any skill, takes some hard work, and lots of practice. But nowhere does it seem to be as hard to acquire as it is for so many of our children in school. The point is that just about anybody can learn to do it if they are organized to put in the time (Singer, 1977). Just about anybody can learn how to read if the organizational formats available to them allow them to spend a sufficient amount of time on task. We must give more attention to how we might help to organize the circumstances for more reading to take place (Singer & Beardsley, 1970; Smith, 1976). Then if we find that it is important to have a better grasp of the psychological skills a child needs to develop in order to learn to read, we can do so. At the same time, psychological theories of reading will have to be expanded to account for the differential consequences of particular learning environments.

A third reason for focusing reading research on how children do and do not learn to read in school is political. Such research would allow us to unpack what is apparently a key fact about the organization of differential levels of reading competence in America, namely, that more members of some groups rather than other groups achieve literacy, despite equal numbers of years in school. This is a political football we could all do without, and it would be good to understand how this fact is arranged. Then we can move beyond it to more relevant concerns for individual children.

Traditionally, the approach to the question of differential school performance by members of different social groups has been to look to the characteristics of the different groups for an explanation. There are a number of reasons for rejecting this approach. Whether the groups are described along genetic, cultural, dispositional, or communicative grounds, there are many reasons for describing these explanations, even in their social science guise, as rooted in unexamined local prejudices (Cole, Gay, Glick & Sharp, 1971; D'Andrade, 1973; McDermott & Gospodinoff, in press). But there is an interesting area of research here. If we look into the classroom contexts in which these facts (race, culture, and communicative codes) become salient in the organization of learning behavior, then we may discover also how differential reading skills become salient points of identity and sources of group borders in any given classroom.

Research into the organization of learning to read in one first grade classroom in a successful middle class school has revealed the following scenario (McDermott, 1976, 1977; McDermott & Aron, 1978). Children who entered school lacking reading skills were identified quickly by their teachers and handled differently. The organizational consequence of this differential handling was that they received only one-third the amount of time on reading tasks as the children classified as better readers. Thus, for every day that they were in the classroom, they fell further behind their peers in the development of reading skills.

This scenario was worked out despite the fact that the teacher spent an equal amount of time with all the children, worried deeply about those who were behind, and was one of the more successful teachers with such children. The following fact points to the different experiences of children who were behind. During their time at the reading table with their teacher, the children in the bottom group were interrupted some 40 times as compared to the children in the advanced group who were interrupted only twice. Now consider that two-thirds of the interruptions came from the children from the advanced group who entered the work space of the bottom group or from the teacher who had to stop instruction in order to discipline children in other parts of the classroom. At first, it appears to be a plot against the possibility of the children who are behind ever catching up, and on the level of their learning to read, it just might as well be. But it is not a plot. It is simply everyone trying to do their best given the way the classrooms in America are organized. The problem with our schools is not that children and teachers do not try hard, but that trying hard does not help to solve the organizational dilemmas which flow from pitting children against each other on the basis of their reading skills.

Children who come into school behind in the development of reading skills are treated differently, and often unfairly, because they represent pedagogical and interactional problems for their teacher. Most often, their lack of skills has nothing to do with their ability to learn to read; often it means that their parents have not taken the time to teach them to read at home, or that they are natives to another language or culture. Or you can claim that they are less developmentally advanced than the better readers, or, if you must, but I will not join you, that they are not as innately smart as their peers. The reason for their being behind at age six need not make a permanent difference. They can still learn how to read at seven, eight or ten (Downing, 1973; Rohwer, 1971). But once they get only one-third the amount of time on task as their peers, they are in trouble, so much trouble that with time, they give up on being unfavorably compared to their peers on reading and begin to develop other identities such as class clown, bully, flirt, rebel (Labov & Robins, 1969; Piestrup, 1973; Shuy, 1977).

Children who come in behind are a pedagogical problem, because they need more of the teacher's time to catch up. They are an interactional problem, because being behind is an embarrassment which disrupts their relations with the other children and the teacher. *The question in American classrooms is not whether or not everyone can learn to read, but how fast a child can learn to read and how much better than others in the same class.* Identities and institutional biographies are on the line. Everyone is sorting everyone else out into the successful and unsuccessful piles. To make matters worse, everyone is doing this while making believe that this is not what is happening. The handling of these identities and coverups takes time and generates the organizational mishaps which keep the problem children from spending time on task, from spending more time on becoming less of a problem. Despite our ideology of equal opportunity for all, our classrooms are organized around sorting children into those who can make it and those who cannot. Children who bring any differences to the classroom are easy marks for being sorted out. However, if the natural variation in culture, language, interest and skill was not available, no doubt we would have to make some even more arbitrary ways of distinguishing the good from the bad, the successful from the problem, the smart from the dumb.

If this scenario has any relation to how schools try to develop reading skills in children, then much of the research presently getting done on reading represents a limited effort which keeps us from understanding the real problem of the reading problems in our schools. The real problem is that our schools appear to be organized, no matter how unconsciously, no matter how much by hard working, caring and devoted people, to produce a certain percentage of school failures among any new generation of school children. This is a social organizational problem and demands a social organizational solution. Neither research nor training in the skills underlying reading competence will help us alter the social situation. Work on motivating better readers will have to start with this problem.

Acknowledgments

I greatly appreciate Lois Hood, Sylvia Scribner, Stanley Wanat and two anonymous reviewers for this Yearbook, all of whom offered insightful suggestions on the original draft of this paper.

References

Bronfenbrenner, U. Toward an experimental ecology of human development. *American Psychologist,* 1977, *32,* 513-531.

Cazden, C. Learning to read in classroom interaction. In L. Resnick & P. Weaver (Eds.), *Theory and practice in early reading.* Hillsdale, NJ: Lawrence Erlbaum, in press.

Cole, M., Gay, J., Glick, J. & Sharp, D. *The cultural context of learning and thinking.* New York: Basic Books, 1971.

Cole, M., Hood, L. & McDermott, R. *Ecological niche picking: Ecological invalidity as an axiom of experimental cognitive psychology.* Laboratory of Comparative Human Cognition and the Institute of Comparative Human Development, Working Paper 14. The Rockefeller University, 1978.

Cole, M., Sharp, D. & Lave, C. The cognitive consequences of education. *The Urban Review,* 1976, *9,* 218-233.

Conklin, H. Bamboo literacy in Mindoro. *Pacific Discovery,* 1949, 2, 4-11.

D'Andrade, R. Cultural constructions of reality. In L. Nader & T. Maretzki (Eds.), *Cultural illness and health.* Washington, DC: American Anthropological Association, 1973.

Downing, J. *Comparative reading.* New York: Macmillan, 1973.

Estes, W. The locus of inferential and perceptual processes in letter identification. *Journal of Experimental Psychology: General,* 1975, *104,* 122-145.

Fredericksen, C. Discourse comprehension and early reading. In L. Resnick & P. Weaver (Eds.), *Theory and practice in early reading.* Hillsdale, NJ: Lawrence Erlbaum, in press.

Griffin, P. Reading and pragmatics: symbiosis. In R. Shuy (Ed.), *Linguistic theory.* Newark, DL: International Reading Association, 1977.

Labov, W. & Robins, C. A note on the relationship of reading failure to peer-group status in urban ghettos. *Florida FL Reporter,* 1969, *7,* 54-57, 167.

McDermott, R. *Kids make sense: An ethnographic account of the interactional management of success and failure in one first grade classroom.* Unpublished doctoral dissertation, Stanford University, 1976.

McDermott, R. Social relations as contexts for learning in school. *Harvard Educational Review,* 1977, *47,* 198-213.

McDermott, R. & Aron, J. Pirandello in the classroom. In M. Reynolds (Ed.), *The futures of education.* Reston, VA: Council for Exceptional Children, 1978.

McDermott, R. & Gospodinoff, K. Social contexts for ethnic borders and school failure. In A. Wolfgang (Ed.), *Nonverbal behavior.* New York: Academic Press, in press.

Piestrup, A. *Black dialect interference and accommodation of reading instruction in first grade.* Berkeley: Monographs of the Language-Behavior Research Laboratory, 1973.

Rohwer, W. Prime time for education. *Harvard Educational Review,* 1971, *41,* 316-341.

Scribner, S. Situating the experiment in cross-cultural research. In K. Riegel & J. Meacham (Eds.), *The developing individual in a changing world.* The Hague: Mouton, 1978.

Scribner, S. & Cole, M. Unpackaging literacy. *Social Science Information,* 1978, *17,* in press.

Scribner, S. & Cole, M. Literacy without schooling: Testing for intellectual effects. *Harvard Educational Review,* 1978, *48,* in press.

Shuy, R. Sociolinguistics. In R. Shuy (Ed.), *Linguistic theory.* Newark, DL: International Reading Association, 1977.

Singer, H. IQ is and is not related to reading. In S. Wanat (Ed.), *Issues in evaluating reading.* Arlington, VA: Center for Applied Linguistics, 1977.

Singer, H. & Beardsley, B. Motivating a disabled reader. *Thirty-seventh yearbook of the Claremont College Reading Conference.* Claremont, CA: Claremont College, 1970.

Smith, D. *A technology of reading and writing* (Vols. 1 & 3). New York, NY: Academic Press, 1976.

Smith, E. & Kleiman, G. Word perception: Theoretical issues and instructional hints. In L. Resnick & P. Weaver (Eds.), *Theory and practice in early reading.* Hillsdale, NJ: Lawrence Erlbaum, in press.

Soderbergh, R. *Project early reading* (Report No. 2). Stockholm University, 1973.

Sticht, T. (Ed.), *Reading for working*. Washington: HumRRO, 1975.

Szwed, J. The ethnography of literacy. Paper presented to the NIE Conference on Writing, Los Angeles, 1977.

Wanat, S. Language behind the eye. In R. Monty & J. Senders (Eds.), *Eye movements and psychological processing*. Hillsdale, NJ: Lawrence Erlbaum, 1976.

THOMAS W. BEAN, PH.D.
California State University
Northridge, California

The influence of a structured overview on comprehension and oral reading miscues of selected college students

The influence of prior knowledge on the meaningful reception, integration, and retention of new concepts is a widely accepted precept in learning theory (Ausubel, 1968; Smith, 1975). A group of reading professors asked to read a passage on ocean engineering will display precise, yet meaningless word calling (Goodman, 1977). This is hardly surprising, since the requisite nonvisual information available to the practicing ocean engineer is simply nonexistent for the reading professor.

An absence of prior knowledge about a particular subject, compounded by a lack of teacher guidance in developing cognitive readiness may impede a learner's understanding of new concepts. This is not a unique or earthshattering notion, yet many university and community college reading improvement programs continue to "plug" students into pre-packaged commercial kits and collections of articles with scant attention to cognitive readiness (Karlin, 1977).

A number of studies have been conducted in content area classes to examine the efficacy of pre-teaching concepts and key vocabulary with a graphic advance organizer or structured overview (Vacca, 1977). A structured overview presents the learner with a hierarchical map for pre-reading discussion aimed at integrating his prior knowledge with new information. Earle (1973) reviewed research in cognitive readiness and recommended that structured overviews receive additional research attention in the content area classroom.

The purpose of the two exploratory studies reported here was to examine the effect of a structured overview on the comprehension and oral reading miscues of students enrolled in a college reading improvement course. At this stage of investigation, the focus was on identifying trends in the data that might suggest areas for more detailed study with a population of students outside the content area classroom.

Silent Reading Study

The first study examined the dependent variable of silent reading comprehension under three conditions. Specifically, would there be any significant difference in comprehension for students receiving pre-reading instruction with a relevant structured overview and students receiving either a placebo overview on an unrelated topic, or no pre-reading instruction?

Method

Eighty-five university students enrolled in a reading improvement course participated in either the structured overview treatment or control conditions. Students participated in the study according to their respective instructional periods. Thus, students in period five recieved the structured overview in a

228

teacher guided pre-reading discussion (n = 28), period six, the placebo (n = 27), and period seven, no pre-instruction (n = 30). Subjects in this exploratory study were not randomly assigned to treatment or control conditions, and individual reading achievement levels were not taken into consideration.

Each of the three conditions involved silent reading of a Health Education article (McPherson, 1972) from a collection of reading improvement selections. The article was at the college level of difficulty according to the Fry Graph(1968). Following the reading, students took a 15 item, researcher constructed comprehension test highlighting cause-effect relationships from the selection. No validity or reliability data was computed for this instrument.

Results

Three *t*-test comparisons were made to evaluate differences among the three groups. The difference between the structured overview and control group was statistically significant for silent reading comprehension, $t\ (56) = 3.72$, $p < .05$. A significant difference was also revealed between the structured overview and placebo group, $t\ (53) = 2.40$, $p < .05$. And comparison of the placebo and control group evidenced statistical significance, $t\ (55) = 2.35$, $p < .05$. The means for the structured overview, placebo, and control groups were 14.25, 10.63, and 10.16 on the 15-item comprehension test.

Oral Reading Study

A related study was conducted to examine the effect of the structured overview on the post oral reading comprehension and quality of miscues generated by a sample of university students enrolled in the reading improvement course.

An assumption implicit in the construction and use of the structured overview is that pre-reading exposure to technical vocabulary, coupled with teacher pronunciation of unfamiliar terms, will reduce the amount of nonproductive encoding a student resorts to while reading (Vacca, 1977). If this assumption holds, students should reveal fewer miscues on vocabulary that was pre-taught with the structured overview. Thus, the following three questions were examined:

(1) Would there be any statistically significant difference in the post oral reading comprehension test scores of the structured overview and placebo group?
(2) Would there be any significant difference in the quantity of miscues generated on vocabulary pre-taught with the structured overview in comparison to the placebo group?
(3) Would there be any significant difference in the quantity of semantically unacceptable miscues generated by the structured overview and placebo group?

Method

Ten university students enrolled in a reading improvement course were randomly assigned to either a structured overview condition (n = 5) or placebo overview condition (n = 5). In both conditions students read the Health Education passage orally after receiving pre-instruction with a relevant structured overview or placebo overview. The oral reading was taped and subscribed to the scenario for administration and analysis in the *Reading Miscue Inventory* (Goodman & Burke, 1972). Following the oral reading, students took a four item, researcher constructed short essay test highlighting cause-effect relationships from the selection. No validity or reliability data was computed for this instrument.

Results

Three t-test comparisons were made for each of the questions under consideration. The difference between the structured overview group and placebo group on comprehension was not statistically significant, $t(8) = 1.90$, $p > .05$, but did reach slight significance at $p < .10$. A difference in quantity of miscues generated on vocabulary pre-taught was not detected, $t(8) = -1.44$, $p > .05$; $p > .10$. The quantity of semantically unacceptable miscues did not reach statistical significance for the two groups, $t(8) = .867$, $p > .05$; $p > .10$.

Discussion

The results of the first study seem to offer tentative support for the structured overview as a means of guiding and increasing silent reading comprehension in the reading improvement classroom. Even a placebo overview seems to develop in students a mind set to approach a reading selection with a predisposition for meaningful retention of major concepts and subordinate ideas.

This exploratory study needs to be replicated with some modifications. That is, subjects should be randomly assigned to treatment and control conditions. Independent variables such as reading achievement, validity and reliability of the comprehension test, and prose characteristics of the reading selection might be considered. Other dependent variables such as delayed recall might provide information on the degree to which a structured overview functions as a mnemonic device.

The second exploratory study required subjects to read the selection orally, an unnatural task at secondary level. This feature, in conjunction with the small sample size, provided little information on the degree to which pre-teaching vocabulary with a structured overview facilitates fluent reading. Perhaps a more refined version of this study would produce more useful data. The procedures used in this study would require a fairly large staff (e.g., 10) of trained miscue examiners to collect the oral reading samples.

In summary, the first study reported here lends support to Earle's (1973) contention that the structured overview is both a valuable means of providing cognitive readiness, and a device warranting continued exploration in a variety of classroom settings.

References

Ausubel, D. P. *Educational psychology — A cognitive view.* New York, N.Y.: Holt, Rinehart and Winston, 1968.

Earle, R. Cognitive readiness: An overview of research. In H. Herber and R. Barron (Eds.), *Research in reading in the content areas: Second year report.* N.Y.: Syracuse University Reading and Language Arts Center, 1973.

Fry, E. The readability formula that saves time. *Journal of Reading,* 1968, *11,* 513-516, 575-578.

Goodman, K. S. *Reading comprehension.* Presentation at a workshop sponsored by the Center for the Study of Reading, University of Illinois, Urbana, Spring 1977.

Goodman, Y. M. & Burke, C. L. *Reading Miscue Inventory.* New York: Macmillan Co., Inc., 1972.

Karlin, R. *Teaching reading in high school: Improving reading in content areas.* Indianapolis: Bobbs-Merrill Co., Inc., 1977.

McPherson, J. Killers in the air. In L. Miller (Ed.), *Developing reading efficiency.* Minneapolis: Burgess Publishing Co., 1972.

Smith, F. *Comprehension and learning.* New York, N.Y.: Holt, Rinehart and Winston, 1975.

Vacca, R. T. Readiness to read content area assignments. *Journal of Reading,* 1977, 20, 387-392.

RICHARD T. VACCA
The University of Connecticut

The effect of expanded directions and adjunct aids on students' comprehension of world history text

There has been substantial research reported on the cognitive processing of readers in prose learning situations. Experimental emphasis in prose learning research, by and large, has been on what the reader does during the learning episode. Wittrock (1963) indicates, "The result of guidance in learning must be evaluated in terms of what the guidance causes the student to do" (p. 74). Emphasis on what the reader does, according to Faw and Waller (1976), has led to a "simple" methodology in the study of prose learning: "manipulate the students' activities during acquisition by means of special instructions, suggestions, or questions, and observe what effects these variations have on learning and retention" (p 692). The problem, then, has been one of determining how experimental manipulation modifies cognitive processing in a way that facilitates what will be learned.

Manipulation, in general, has either been preparatory in nature, i.e., advance organizers (Ausubel, 1963), structured overviews (Barron, 1971), and directions (Wittrock, 1963), in order to modify the reader's set to learn, or else has been guiding in nature, i.e., inserted questions (Frase, 1970, 1977), and reading guides (Herber & Vacca, 1977), in order to modify the ongoing processing behaviors of the reader. Frase (1971), in fact has proposed a heuristic model for research in prose learning in which he delineates the role of adjunct aids to account for the utility of any class of verbal stimuli (advance organizers, directions, questions, statements, etc.) in facilitating learning from text. Adjunct aids may be thought of as "orienting directions" (Frase, 1970) which dispose the reader to respond actively to certain aspects of text.

Questions doubtlessly form a subclass of adjunct aids which has received considerable attention from reading researchers (Smith, 1977). Three characteristics of questions appear to have facilitative effects on text learning: 1) the position of questions in text, 2) the contiguity of questions to related content, and 3) the type of question asked. It is assumed by this investigator that these characteristics hold across various classes of adjunct aids other than questions.

Rickards (1975-76), for example, has examined the process effects of advance organizers interspersed in text. He varied the type of organizing statements (superordinate and coordinate) and their position either before (prestatement) or after (poststatement) associated segments of text. Through such experimental manipulation, Rickards determined that superordinate prestatements (advance organizers) not only led to significantly greater contiguous recall of subordinate facts but also generated recall of information derivable from paragraphs incidental to the organizers themselves.

In the present study the investigator examined the facilitative effect created by expanded directions and adjunct aids on students' comprehenion of world history text in tenth grade social studies. It was predicted that a preparatory set to learn, induced through expanded directions (Barron, 1971), and adjunct aids positioned *en masse* either before or after the text selection, or interspersed individually within the text selection, would have a positive effect on students' comprehension of world history content.

Method

Subjects

The students were drawn from the tenth grade of a parochial high school located in a Northeastern city. Although an entrance test is required for admittance, subjects represented a wide range of ability levels. The students, on the whole, came from lower middle to upper middle socio-economic level homes. Nearly all participants were Caucasian.

Ninety-seven pupils comprised the original sample. Only 74 students, however, were included in the final data analysis. In order to have been included in the final sample, students must have participated in every aspect of the experimental treatment.

Two teachers participated in the experiment. One teacher was responsible for three treatment groups and the other for two treatment groups.

Procedure

Five intact world history classes were assigned randomly to one of five treatments. Below is a description of the treatment groups:

Experimental Group 1 (E_1):	Received expanded directions plus pre-exposure to adjunct aids for each text selection
Experimental Group 2 (E_2):	Received expanded directions plus post-exposure to adjunct aids for each text selection
Experimental Group 3 (E_3):	Received expanded directions plus adjunct aids interspersed within each text selection
Experimental Group 4 (E_4):	Received expanded directions plus required to read each text selection without any exposure to adjunct aids
Control Group 1 (C_1):	Required to read each text selection without exposure to expanded directions or adjunct aids

Each of the experimental groups was exposed to a two day period of expanded directions as specified by Barron (1971). Expanded directions were operationalized to include: 1) the reading of the prepared handout on the importance of perceiving thought-relationships in text and drawing interpretations from them, 2) a review of the basic patterns of thought-relationships (cause-effect, comparison-contrast, time order, enumeration) and the "reading signals" which help identify these patterns, 3) completion of exercises to provide students with practice in identifying the patterns as they appear in sentences and paragraphs taken from the textbook, and 4) guided explanation of the use of adjunct aids in each of their respective treatment groups.

The adjunct aids for each text selection were comprised of a set of statements which represented interpretations of the text selection. A set of statements was positioned *en masse* either before or after each selection of text in groups E_1 and E_2 respectively. In group E_3 each adjunct statement was individually embedded immediately after that segment of the selection which revealed the intratext relationship that led to an interpretation. If the students in groups E_1, E_2 or E_3 believed a statement to be a reasonable or "correct" interpretation of the text, they checked the appropriate space to the left of each statement. This procedure is recommended by Herber (1970) as a way of maintaining the reader's active response to each statement.

Group E_4 received expanded directions as operationalized above minus step 4 (the explanation of how to respond to adjunct aids). The control group received two study hall periods in place of expanded directions. All groups were given 45 minutes to complete each text selection, with or without adjunct aids.

Three sections from two chapters in a world history textbook were selected for reading. The selections, "The Paris Peace Conference," "The Spread of Democracy," and "The Growth of Nationalism," were chosen, essentially, for their length (1000-1500 words) and their relevance to the unit tipic. "Post World War One Era." The investigator developed the set of adjunct aids which accompanied each reading.

The experiment took six school days to complete. Monday and Tuesday of the first week were devoted to expanded directions. Students were then given successive text selections on Wednesday, Thursday and Friday. On Monday of the second week, all groups were administerred a 30-item multiple choice test in which 45 percent of the questions tested knowledge, 45 percent tested comprehension and 10 percent tested application as outlined in *The Taxonomy of Educational Objectives* (Bloom, 1956).

Design and Statistical Analysis

The basic design of the study was the "Non-Equivalent Control Group Design" (Campbell and Stanley, 1963), since the experimenter had only intact classes at his disposal.

Analysis of covariance was used to analyze the data. The subtest, "Social Studies Reading," from *S.R.A. National Education Development Tests* was the covariate used to yield scores that controlled for the criterion variable, scores on the world history content test. If the overall F (with alpha level set at .05) indicated significant differences among treatments for comprehension of social studies content, all pairwise comparisons of means were to be evaluated by a post hoc technique developed by Tukey.

Results

The hypothesis, cast in its null form for statistical analysis, stated that there would be no differences in ability to comprehend world history text among students in any of the five treatment groups. The analysis of covariance, however, indicated a significant overall difference among the groups, $F(4,68) = 5.80$, $p < .05$, thus rejecting the null hypothesis. The means for the five groups on the 30-item comprehension test were 21.14, 18.72, 17.75, 16.60 and 16.58.

Post hoc inspection of the data determined that students in E_3 (expanded directions combined with interspersed adjunct aids) comprehended the text selections significantly better than E_4 (expanded directions and reading-only) and C_1 (reading only).

Discussion

The observed outcome of the study, that students receiving expanded directions combined with interspersed adjunct aids (E_3) performed significantly better than either students receiving expanded directions and reading-only (E_4) or reading only (C_1) may be attributed to several factors. With respect to positioning adjunct aids in text, the study reaffirms a consistent observation in prose learning research. That is, adjunct aids "work" best when they occur immediately after relevant portions of text (Frase, 1970; Faw

& Waller, 1976). Whereas pre-questions function to cue responses to specific stimuli, interspersed questions "placed after the related text . . . reinforce learning behaviors which are not limited to specific stimuli" (Frase, 1970, p. 345). Since the outcome measure for the study contained questions at the knowledge, comprehension and application levels of Bloom's taxonomy, the E_3 treatment obstensibly exerted influence on cognitive processing beyond intentional or statement-specific information and, indeed, had a facilitative effect on total learning.

It should be noted, further, that the statements in experimental conditions E_1, E_2, and E_3 were superordinate in nature. As described by Rickards (1975-76), a superordinate statement is defined as "a single sentence representing a topic or overriding concept implied" in a text segment, i.e., an interpretation. Thus the adjunct aids positioned immediately after relevant text segments may have had a post organizer effect which facilitated the search activities of the students in E_3.

This study also suggests that provisions for a process set alone may be too generalized to guide learning. Simply knowing that a particular behavior should be used in a learning situation, and establishing its importance to content acqusition, is not sufficient. Probably students who lack maturity as learners require structured guidance in ongoing processing activities during the reading act.

References

Ausubel, D. P. *Educational psychology, a cognitive view.* New York: Holt, Rinehart and Winston, Inc., 1968.

Barron, R. F. *An iterative-ecological investigation of vocabulary development in tenth grade biology.* Unpublished doctoral dissertation, Syracuse University, 1971.

Bloom, B. S. *Taxonomy of educational objectives — Handbook I: Cognitive domain.* New York: David McKay, 1956.

Campbell, D. T. & Stanley, J. C. *Experimental and quasi-experimental designs for research.* Chicago, Illinois: Rand-McNally and Company, 1963.

Faw, H. W. & Waller, T. G. Mathemagenic behaviors and efficiency in learning from prose. *Review of Educational Research,* 1976, 691-720.

Frase, L. T. *A heuristic model for research on prose learning.* Paper presented at the Annual Meeting of the American Education Research Association, New York, 1971.

Frase, L. T. Boundary conditions for mathemagenic behaviors. *Review of Educational Research,* 1970, 337-347.

Frase, L. T. Purpose in reading. In J. T. Guthrie (Ed.), *Cognition, curriculum and comprehension.* Newark, Delaware: IRA, 1977.

Herber, H. L. & Vacca, R. T. (Eds.). *Research in reading in the content areas: Third report.* Syracuse, New York: Reading and Language Arts Center, Syracuse University, 1977.

Herber, H. L. *Teaching reading in content areas.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970.

Rickards, J. P. Processing effects of advance organizers interspersed in text. *Reading Research Quarterly,* 1975-76, 599-622.

Smith, A. E. Pre-reading questions as purposes for reading. In H. L. Herber & R. T. Vacca (Eds.), *Research in reading in the content areas: Third report.* Syracuse, New York: Reading and Language Arts Center, Syracuse University, 1977.

Watts, G. H. & Anderson, R. C. Effects of three types of inserted questions on learning from prose. *Journal of Educational Psychology,* 1971, 387-394.

Wittrock, M. C. Effects of certain sets upon complex verbal learning. *Journal of Educational Psychology,* 1963, 85-88.

ANNE E. WOLF
University of Wisconsin-Eau Claire

LARRY MIKULECKY
Indiana University-Bloomington

Effects of uninterrupted sustained silent reading and of reading skills instruction on changes in secondary students' reading attitudes and achievement

Lyman Hunt (1970) introduced the concept of Uninterrupted Sustained Silent Reading, or USSR. Proponents of USSR, such as Allington (1975), Ganz and Theobold (1974), McCracken (1971), and Mork (1972), have strongly recommended adoption of the program, which allows students to read self-selected materials in school on a regularly scheduled basis. USSR programs may also be referred to as SSR (Sustained Silent Reading), HIP (High Intensity Practice), Reading Break, or Free Reading.

Research on the effectiveness of USSR is limited and inconsistent. On the elementary level, USSR, in conjunction with some skills training, has been found to be as effective in skills improvement as an equivalent amount of time spent in skills-only instruction (Evans & Towner, 1975; Hanson, 1972; Oliver 1970). Hanson also found that USSR had a positive effect on reading habits. Wilmot (1975) found significant improvement in reading attitudes among USSR groups of elementary students, but control groups with no USSR improved significantly more in reading comprehension; no significant differences were found in vocabulary growth. Petre (1971) provided several anecdotal reports of successes in 50-75 Reading Break programs in Maryland elementary schools.

Little research has been done on the effects of USSR on older students. Mikulecky and Wolf (1977) reported no significant differences among seventh grade subjects experiencing USSR and two other treatments in required reading classes. Attitude toward reading declined in all groups, although the USSR group had the least decline. The overall decline was interpreted as part of the general phenomenon of decline in attitude toward reading as more time is spent in school (Bullen, 1972; Mikulecky, 1976).

Although USSR can be implemented in any classroom setting, at secondary levels it seems most often attempted as part of required reading courses. This study examined the effects of two typical treatments, Skills Instruction Alone and Skills Instruction Plus USSR, on the reading achievement and reading attitudes of seventh grade students in required nine-week developmental reading classes. A single research question was investigated: Did students who experienced 100% skills instruction differ in reading achievement and attitude toward reading from students who experienced 80% skills instruction and 20% USSR?

Method

Subjects

The 125 subjects were white seventh grade students attending junior high school in a suburban community near Milwaukee.

Procedure

There were two treatments: Skills Instruction Alone and Skills Instruction Plus USSR. Subjects experiencing Skills Instruction Alone worked for five 40-minute periods a week for nine weeks on reading skills units such as context clues, word structure, and main idea. Subjects experiencing Skills Instruction Plus USSR worked for four 40-minute periods a week on the same skills units taught in the skills-alone treatment. However, for one 40-minute period each week subjects read self-selected materials brought from home or chosen from a classroom library.

Treatments were conducted simultaneously for the duration of a nine-week required developmental reading course. Two classes were randomly assigned to Skills Instruction Alone and three classes to Skills Instruction Plus USSR. Subjects were pretested with two measures of reading achievement, the vocabulary and comprehension subtests of the *Stanford Diagnostic Reading Test* (Level II, Form W), and with two measures of attitude toward reading, the Estes scale (Estes, 1971) and the Mikulecky Behavioral Reading Attitude Measure (MBRAM) (Mikulecky, 1976). Students were posttested with parallel forms of the achievement measures and with the single forms of the attitude measures.

Data Analysis

One-way univariate analyses of covariance were calculated for the comprehension, vocabulary, and attitude posttest scores with the corresponding pretest score as covariate.

Results

Analysis of covariance for vocabulary showed no significant differences between the two treatment groups, F (1,102) = .07, $p > .05$. Mean raw scores for both groups declined from the pretest to the posttest. However, the decline was less than one point for either group which was well within the subtest's standard error of measurement of 2.2 points. Analysis of covariance for comprehension showed a significant difference between treatment groups, F (1,100) = 4.61, $p < .05$. Mean raw scores declined for both the Skills Instruction Alone group (−3.65 points from 48.72 to 45.07) and the Skills Instruction Plus USSR group (−1.03 points from 47.97 to 46.94). While the decline was significantly smaller in the group experiencing both skills instruction and USSR, both declines were close to or within the subtest's reported standard error of measurement of 3.0 points.

Analysis of covariance for both the MBRAM and the Estes scale indicated no significant differences in attitude toward reading between the two treatment groups, F (1,104) = 2.55, $p > .05$ and F (1,101) = 1.67, $p > .05$. Within each treatment group, mean raw scores declined from the pretest to the posttest. For the MBRAM these nonsignificant declines were −.89 (60.13 to 59.24) for Skills Instruction Alone and −5.07 (60.77 to 55.70) for Skills Instruction Plus USSR, and for the Estes scale −2.92 (71.62 to 68.70) and −4.37 (74.25 to 69.88) respectively. There was a greater, but not significantly greater, decline in attitude in the group experiencing USSR.

Lack of significant difference between treatment groups on the vocabulary measure and a significantly smaller decline on the comprehension measure for the Skills Instruction Plus USSR group indicate that 20% more time in skills instruction did not benefit students' reading achievement. Also, one day a week spent in USSR did not result in either a gain or a smaller decline in attitude for the Skills Instruction Plus USSR group.

In both treatment groups, raw scores declined on all measures from pretest to posttest. Declines on the achievement measures probably resulted from standard error of measurement or from a ceiling effect, since better readers had scored very well on the pretest with little room for growth on the posttest. The declines are probably not educationally significant. Declines on the reading attitude measures ranged from less than two points to a high of over ten points. Only one class in either treatment group experienced a decline of more than two points on both measures. This class had expressed great dislike for the course. Attitude declines probably represent the typical decrease in positive attitude experienced as students go through school or resentment at being required to take the reading course.

This study suggests that shortterm reading courses may have little impact on reading achievement scores and may be detrimental to attitude toward reading even when USSR is incorporated into the coursework. USSR need not, however, be linked to a reading course, and research should be conducted on its effects at different grade levels, over extended periods of time, and in different instructional settings.

References

Allington, R. Sustained approaches to reading and writing. *Language Arts*, 1975, *52*, 813-815.

Bullen, G. F. *The development and validation of a reading attitude measure for elementary school children*. Washington, D. C.: U.S. Office of Education Project No. 1A045, 1972. (ERIC Document Reproduction Service No. ED 109 211)

Estes, T. H. A scale to measure attitudes toward reading. *Journal of Reading*, 1971, *15*, 135-138.

Evans, H. M. & Towner, J. C. Sustained silent reading: Does it increase skills? *Reading Teacher*, 1975, *29*, 155-156.

Ganz, P. & Theobold, M. B. Suggestions for starting SSR. *Journal of Reading*, 1974, *17*, 614-616.

Hanson, R. A. A study of the effects of SRR on the reading attitudes and habits of second, third, and fourth grade students in a low income area (Doctoral dissertation, University of North Dakota, 1972). *Dissertation Abstracts International*, 1973, *33*, 5034A. (University Microfilms No. 73-5965.).

Hunt, L. The effect of self-selection, interest and motivation upon independent, instructional, and frustrational levels. *Reading Teacher*, 1970, *24*, 146-151.

Karlsen, B., Madden, R. & Gardner, E. F. *Stanford Diagnostic Reading Test*, Level II, Forms X and W. New York: Harcourt Brace, 1966.

McCracken, R. A. Initiating sustained silent reading. *Journal of Reading*, 1971, *14*, 521-524, 582-583.

Mikulecky, L. *The developing, field-testing, and initial norming of a secondary/adult level reading attitude measure that is behaviorally oriented and based on Krathwahl's Taxonomy of the Affective Domain*. Unpublished doctoral dissertation, University of Wisconsin-Madison, 1976.

Mikulecky, L. & Wolf, A. The effect of uninterrupted sustained silent reading and of reading games on changes in secondary students' reading attitudes. In P. D. Pearson (Ed.), *Twenty-sixth Yearbook of the National Reading Conference*. Clemson, S. C.: National Reading Conference, 1977.

Mork, T. A. SSR in the classroom. *Reading Teacher*, 1972, *25*, 438-441.

Oliver, M. E. The effect of high intensity practice on reading achievement (Doctoral dissertation, Syracuse University, 1970). *Dissertation Abstracts International*, 1971, *32*, 311A. (University Microfilms No. 71-18,501

Petre, R. M. Reading breaks make it in Maryland, *Journal of Reading*, 1971, *15*, 191-194.

Wilmot, M. P. *An investigation of the effect upon reading performance and attitude toward reading of elementary grade students, of including in the reading program a period of sustained silent reading*. University of Colorado, 1975. (ERIC Document Reproduction Service No. ED 119 130)

MARILYN G. EANET
Rhode Island College

An investigation of the REAP reading/study procedure: its rationale and efficacy

The REAP procedure is a teaching-learning strategy designed to improve reading, writing, and study skills. The focus of this study is upon advancing the design of the strategy and assessing various aspects of its use as a reading/study procedure within the setting of a college reading improvement course.

The REAP Reading/Study Procedure has four basic steps. The student is asked to:

R — *Read* to discover the writer's message.
E — *Encode* the message by putting it into his/her own language.
A — Annodate by cogently writing the message in notes for him/herself or in a thought book to share with others.
P —*Ponder*, in other words, process the message, now in annotation form, through internal "dialectic thinking" or through discussion with others.

Central to the REAP procedure is having students learn to write annotations specially designed to achieve certain learning objectives (Manzo, 1973). In writing such annotations, the reader is required to discriminate and synthesize the ideas presented by the writer, translate these into his/her own language, and crystallize the result in writing.

Seven annotation forms have been developed in working with the procedure: summary, thesis, question, critical, heuristic, intention, and motivation annotations. Some annotations are more suitable to some types of writing than are others (Eanet & Manzo, 1976). The present study with its emphasis on study reading utilized the first four most frequently.

The theoretical rationale for the REAP procedure finds a supportive base in several related areas of research in reading and learning.

Gagne (1973), among others, has hypothesized that presenting a learner with a specific learning task (for example, the charge to annotate) activates a "set" within him that determines the "strategies" with which he approaches the material he is to read. He concludes that these strategies of "search, selecting, and coding" in large measure determine what the learner will retain.

Supporting evidence for the existence of these strategies and their flexibility is plentiful. Anderson and Biddle (1975), in their comprehensive review of adjunct questions, find the explanation that such questions effect performance "by influencing the processing activities of readers" most persuasive. A number of prose learning studies give evidence that other task or "set" variables can also determine what a reader remembers, (e.g., Frase, 1975).

The use of an overt, constructed response as a productive learning task finds some support in studies of adjunct questions (Anderson & Biddle, 1975) and programmed learning (Williams, 1963, 1965), although in the studies reviewed, the responses were not as complex as are the REAP annotation forms. A further note of caution came from early investigations of study-reading which involved precise or summary writing and failed to show support for such an approach (Germane, 1921; Mathews, 1938; Arnold, 1942; Stordahl & Christensen, 1956).

A major question arising from close examination of many related studies is whether the learning tasks and the strategies they induce actually reflect an increase in the *quality* of study or whether their effectiveness lies in their ability to increase *quantity* of study by lengthening the time spent inspecting and reflecting upon the material to be learned. In the absence of persuasive evidence either way, inspection time was controlled in the present study.

An additional area of interest that arose from a pilot study was how learning the REAP procedure might interact with a student's cognitive style. Since it appears to require a degree of reflection not typical of impulsive learners, it was hypothesized that they might most benefit from its use.

This study, then, focuses upon the teaching and use of a reading/study procedure in which the central feature is a learning task, viz., writing annotations, presumed to have a positive effect upon the reader's learning strategies and thus upon his reading effectiveness. It was conducted as a teaching experiment in a natural setting and centered upon three major issues: 1) the experimental subjects' degree of success in learning and using the procedure, 2) how such learning interacted with the subjects' tendency toward a reflective or impulsive cognitive style, and 3) how subjects trained in the REAP procedure compared with control group subjects on several measures of learning from specific prose passages.

Method

Participants in the study were 105 students enrolled in a college reading and study skills class at the University of Missouri-Kansas City. Each of the six sections of the class were randomly assigned to one of three treatment conditions. The experimental group received instruction and practice in the use of the REAP procedure. The experimental-control group received instruction and practice in using the SQ3R method and related skills practice. The control group received instruction centered upon developing "affective readiness" for reading/study skills improvement followed by instruction and practice in study skills areas not directly related to study-reading or comprehension improvement. Each group received four weeks of instruction with the experimenter serving as instructor.

During the first week of the semester, all groups received the Nelson-Denny Reading Test and the Reading Practices Mini-Inventory, an informal measure of reading/study habits. From these measures, it was determined that the groups did not have statistically significant differences in either basic reading achievement or in reading/study habits.

Within the experimental group, measures were taken to determine the students' progress in mastering and using the REAP procedure. These students were also administered the adult version of the Matching Familiar Figures Test, as used by Yando and Kagan (1968), to determine their tendencies toward impulsive or reflective behavior. This cognitive style variable was examined for its relationship to the use and learning of the procedure.

A two by three factorial design with repeated measures on the first factor was employed to examine the experimental subjects' ability to write annotations before and after instruction across the three levels of the cognitive style dimension — impulsive, undetermined, and reflective.

During the sixth week, six posttest measures were administered. Two measures involved written recall of significant ideas in an informative article 1) immediately following a fifteen minute reading/study period and 2) forty-eight hours later. Two objective recognition tests, based upon the same selection, were administered following the delayed recall test. Then, a short philosophical essay was used to elicit two written reflective responses from each subject. All written responses were coded, randomized, typed, and judged against previously established criteria.

Finally, all students responded anonymously to a self-assessment questionnaire designed to determine how useful they perceived their instruction to be and how frequently they were using what they had been taught.

A posttest only control group design was used for the comparative portion of the study.

Results

A two-way analysis of variance for repeated measures was performed on the annotation writing pretest and posttest scores with the cognitive style groups serving as the second factor. Scores on the post measure were significantly higher than those on the pretest, $F(1, 29) = 21.42, p < .001$, with the students obtaining a mean score of 17.63 on the pretest and a mean score of 23.69 on the posttest.

On the second factor, the cognitive style variable, no statistically significant difference was revealed, nor was there a significant interaction. However, an examination of the mean scores for each cognitive style group revealed that the impulsives showed the most improvement in annotation writing with a pretest to posttest difference score of 6.7, while the reflective subjects showed only a 3.9 point improvement.

Data from the six measures designed to compare the experimental (REAP) subjects to the experimental-control (SQ3R), and control subjects in learning in specific reading/study situations were subjected to a one-way analysis of covariance with the Nelson-Denny Reading Test total scores as covariant. The only statistically significant difference was found on the test of immediate recall. A Scheffé multiple comparison procedure applied to the adjusted means determined that the difference lay between the control and the experimental-control groups. Examination of the mean scores on all six measures revealed that, in general, the control group made the highest scores, followed by the experimental (REAP) subjects, and then the experimental-control (SQ3R) subjects.

No significant differences were found among groups on the student self-assessment questionnaire. All groups tended to "Agree Somewhat" that the instruction they received had been helpful. Further, they reported using what they had been taught either "sometimes" or "frequently."

Conclusions and Discussion

College reading improvement students were successful in learning to write the annotation forms central to the REAP procedure. Further, the experimental subjects, as indicated by self-report, found the instruction in the REAP procedure helpful and tended to incorporate it into their study as frequently as their peers did more tradition instruction.

Further, based on an examination of the mean scores, there was some indication that students who demonstrate an impulsive cognitive style may profit most from instruction in the procedure, perhaps because it requires a degree of reflection that they do not typically bring to reading tasks.

However, this study fails to provide evidence that the learning and use of the procedure helped students achieve more effective learning from prose when compared to students who received other training and when reading/study time was controlled. Several possible explanations for these results must be considered. First, it is possible that the learning tasks implicit within the REAP procedure are counterproductive. Secondly, the higher mean scores obtained by the control group may indicate that students had not integrated either REAP or SQ3R into their own reading habits thoroughly enough to compete with other students who studied in their habitual manner. Finally, the time controls may have been too restrictive; the perceived helpfulness of the procedure may lie, not in any tendency to increase quality of study, but rather in increasing inspection and reflection time.

References

Anderson, R. C., & Biddle, W. B. On asking people questions about what they are reading. In G. Bower (Ed.), *Psychology of learning and motivation* (Vol. 9). New York: Academic Press, 1975.

Arnold, H. F. The comparative efficiency of certain study techniques in the field of history. *Journal of Educational Psychology*, 1942, *33*, 449-457.

Eanet, M. G., & Manzo, A. V. REAP — strategy for improving reading/writing/study skills. *Journal of Reading*, 1976, *19*, 647-652.

Frase, L. T. Prose processing. In G. Bower (Ed.), *Psychology of learning and motivation* (Vol. 9). New York: Academic Press, 1975.

Gagne, R. M. Learning and instructional sequence. In F. R. Kerlinger (Ed.), *Review of research in education*. Itasca, Illinois: F. E. Peacock Publishers, Inc., 1973.

Germane, C. E. The value of the written paragraph summary. *Journal of Educational Research*, 1921, *3*, 116-123.

Manzo, A. V. CONPASS English: A demonstration proje[illegible]. *Journal of Reading*, 1973, *16*, 539-545.

Mathews, C. O. Comparison of methods of study for immediate and delayed recall. *Journal of Educational Psychology*, 1938, *29*, 101-106.

Stordahl, K. E., & Christensen, C. M. The effect of study techniques on comprehension and retention. *Journal of Educational Research*, 1956, *49*, 561-570.

Williams, J. P. Comparison of several response modes in a review program. *Journal of Educational Psychology*, 1963, *54*, 253-260.

Williams, J. P. Effectiveness of constructed response and multiple-choice programming modes as a function of test mode. *Journal of Educational Psychology*, 1965, *56*, 111-117.

Yando, R. M. & Kagan, J. The effect of teacher tempo on the child. *Child Development*, 1968, *39*, 27-34.

MARTHA RAPP HAGGARD
Northern Illinois University

The effect of creative thinking - reading activities (CT-RA) on reading comprehension

This study was designed to examine the effect of training in creative thinking on reading achievement. Central to its conception were the following considerations: 1) "creative reading," the acknowledged highest level of reading achievement (Smith, 1969), is apparently more talked about than taught; 2) attempts to teach creative reading might be strengthened through development of systematic teaching procedures; and 3) the Languaging in the Content Areas (LICA) thesis (Manzo & Sherk, 1978) provides a framework for designing procedures which combine creative thinking with reading-language tasks.

"Creative process" is one of the three major notions embodied by the LICA thesis (Manzo & Sherk, 1978). This process, whether called "creative thinking," "divergent thinking," or "creativity," is often perceived as belonging only to those who produce great artistic or literary works. As defined by LICA, however, "creative process," in combination with language learning and dialectical thinking, is a way of looking at the world, i.e., an approach to living, rather than a specific ability or talent. Access to creative process is available to all through instruction, training, and practice.

Within the context of reading, creative process can be viewed as an interaction, or dialogue between reader and text. Just as the creative thinker actively responds to the environment, so too does the creative reader respond to the printed page. This response, according to LICA, takes the form of "dialectical thinking" whereby the reader converses with him/herself, rearranging, adding, and reorganizing information based upon both prior knowledge and printed message.

The Creative Thinking-Reading Activities (CT-RA) used in this study are designed to stimulate dialectical thinking through the inclusion of a language interaction, as suggested by the LICA thesis. Each activity requires that students utilize creative thinking abilities; e.g., improving products, identifying problems, seeking new solutions, etc., then orally retrace their thinking to determine how they arrived at the conclusions, and examine alternative possibilities and/or lines of reasoning. The critical aspects of these games are that 1) they require both divergent and convergent thinking abilities, 2) they include articulation of ideas in the form of a dialogue between teacher and student, and 3) they are engaged in prior to reading instruction so that thought processes used in the activity may be applied to the language-reading task. An example of this type of activity is:

> *CT-RA.* Ask the student to list all of the ways she/he can think of to improve school desks. After the list is completed, discuss each suggestion in terms of how it will make the desk better, how easy or difficult it would be to add, and how important the change would be. Work with the student to assess the likelihood of actually effecting

the changes: what would have to be done? what are the major obstacles? what are the stages for development of changes? should it actually be tried?

Reading-Language Application. After reading a story, choose an important object (significant to the story) and ask the student to list all the ways he/she can think of to improve the object. Enter into a dialogue concerning how these changes might affect the characters, the story outcome, etc.

Method

Sample

The subjects for this study were 33 students enrolled in the 1976 summer reading clinic at the University of Missouri - Kansas City. The students ranged in age from seven to eighteen, and displayed a wide range of reading difficulties.

Procedure

Students received individual instruction from Master's Degree candidates in reading education. The students and their tutors were randomly assigned to one of three treatment groups. Students in the experimental group ($n = 12$) participated in Creative Thinking-Reading Activities, along with tutor-determined remedial instruction. Students in the experimental-control group ($n = 10$) participated in games which are more commonly used in reading/remedial instruction, along with tutor-determined remedial instruction. Students in the control group ($n = 11$) received regular instruction which took the form of tutor-determined procedures for remediation of specific reading problems. The number of students enrolled for the summer clinic fell far below expectancies, thereby sharply reducing anticipated sample sizes. Students and tutors were blind to experimental conditions.

The study was conducted over a period of five weeks. Three days of the first and last weeks were used for testing. During the experimental period, tutors kept daily records of the activities used, along with their own and student reaction to them. The experimenter attended all sessions to observe instructional procedures and to answer to any questions which might arise. Data were collected to answer the following questions:

1. Are there significant differences between the three groups with respect to reading achievement? Are these differences evident when initial achievement levels and IQ are held constant?
2. Are there significant differences between the three groups with respect to creative thinking behavior? Are these differences evident when age and IQ are held constant?

Measures

Individual intelligence test scores (WISC or WAIS) were obtained along with age, sex, and grade data.

The Classroom Reading Inventory (CRI) (Silvaroli, 1973) was used as a measure of reading comprehension. Tutors administered alternate forms of this published informal reading inventory during the first and last weeks of the study.

The Torrance Tests of Creative Thinking, Form A (Torrance, 1966) were used as a posttest only measure of creativity.

211

Results

Reading Achievement

Since the CRI test yields scores from readiness through eighth grade, it was necessary to transform grade-level scores into assigned scores in order to quantify readiness, pre-primer, and primer levels. The highest grade-level score in the instructional range was used for analysis.

To examine differences which existed between the groups with respect to initial reading achievement and IQ, separate analyses of variance were performed on pretest scores and IQ scores. Mean reading pretest scores of 4.7, 6.4, and 4.5 were obtained by the experimental, experimental-control and control groups, respectively. IQ mean scores were 99.6, 102.6, and 99.2. The *F* ratios obtained were not statistically significant.

Using a one-way analysis of covariance, Classroom Reading Inventory posttest scores were adjusted for the effects of initial reading achievement and IQ. Adjusted mean posttest scores for the groups were 6.4, experimental; 5.4, experimental-control; and 6.0, control. A statistically significant treatment effect was found, $F(2,30) = 3.36, p < .05$. The Scheffé multiple comparison procedure was applied to the adjusted mean scores and yielded a statistically significant difference between the experimental and experimental-control groups, $F(2,30) = 3.53, p < .05$. No other statistically significant differences were found.

Creative Thinking

Separate analyses of variance were performed on the verbal, figural and total raw scores from the Torrance Tests of Creative Thinking. No statistically significant differences were found on any of these analyses. Analyses of covariance were then performed, with age and IQ as covariates. Again, no statistically significant differences were found.

Discussion

Interpretation of the Data

The results partially support the hypothesis that training in creative thinking is useful as a means for increasing reading comprehension. Findings from analysis of the Classroom Reading Inventory suggest that students trained in creative thinking benefit in their reading performance from such training. Some support was found for the hypothesis that systematic training in creative thinking would yield differences between the three groups with respect to creativity.

Implications

In general, it seems that the Creative Thinking-Reading Activities, which were designed to act as a warm-up, or catalyst, to creative reading, did increase reading performance. The CT-RA were engaged in prior to the reading lesson itself. They were linked to the reading activity through a process which allowed the student and teacher to review thinking strategies and recapture them while engaged in story-related tasks. Through such a procedure, the reader was then equipped and willing to add to what was being read. This conclusion is strengthened somewhat by the fact that the subjects in this study were acknowledged reading "failures" who exhibited all of the myriad psychological, emotional and educational problems associated with reading disability. That they would exhibit some measure of reading success with the CT-RA serves to emphasize the potency of such an approach.

References

Manzo, A. V. & Sherk, J. K. Reading and languaging in the content areas: A third generational approach. *The New England Reading Association Journal*, 1978, *13*, 28-38.

Silvaroli, N. J. *Classroom Reading Inventory*. Dubuque, IA: Wm. C. Brown Company, 1973.

Smith, N. B. The many faces of reading comprehension. *The Reading Teacher*, 1969, *23*, 249-259.

Torrance, E. P. *Norms-Technical Manual, Torrance Tests of Creative Thinking*. Lexington, Mass.: Ginn & Company, 1974.

Torrance, E. P. *The Torrance Tests of Creative Thinking*. Lexington, Mass.: Personnel Press, 1966.

EUGENE A. JONGSMA
RONALD E. POUND
MARTHA L. TIPS
Southern Methodist University

The effects of instruction in testwiseness in a college reading improvement course

Testing represents the major portion of a college students evaluation and is often the main criterion for admission into graduate and professional schools. The average college student will take over 100 tests during his college career in addition to special examinations such as the *Graduate Record Examination, Law School Admissions Test,* and others. Knowing how to take tests is a skill that is essential to a successful college career. As college students become more and more concerned about their ability to take tests, they are turning to reading and study skills courses for guidance and instruction in test-taking.

Several investigators have examined the effects of teaching testwiseness. Populations have ranged from pre-schoolers to adults. The following conclusions may be drawn from the studies on teaching testwiseness:

1. Instruction in testwiseness is effective when its success is judged on the basis of an internal criterion, usually a project-developed scale of testwiseness (Gibb, 1964; Moore, Schutz, & Baker, 1966; Wahlstrom & Boersma, 1968; Slakter, 1969; Slakter, et al., 1970; Moore, 1971; Pencheff, 1972; Woodley, 1972; Langer, et al., 1973; and Yearby, 1974).

2. Although significant gains in testwiseness are usually found on internal criterion measures as a result of instruction, they frequently do not transfer to external criterion measures (Tinney, 1968; Woodley, 1972; Yearby, 1974; Kahn, 1976; Jongsma, 1975; Keysor & Williams, 1977). Some investigators, however, have found significant differences on standardized achievement and aptitude tests, favoring examinees who have received instruction in testwiseness (Oakland, 1972; Callenbach, 1973; Sause & Grieco, 1973; Gaines & Jongsma, 1974; Eakins, Green, & Bushell, 1976).

3. The long-term effects of instruction in testwiseness have yet to be demonstrated (Callenbach, 1973; Oakland, 1972).

4. There is not enough evidence to suggest that instruction in testwiseness is more beneficial for particular subgroups of students, e.g. disadvantaged students, high ability readers, etc. (Tinney, 1968; Yearby, 1974; Jongsma, 1975).

5. No particular method of instruction (lecture, programmed texts, group discussion) is superior in teaching testwiseness (Langer, et al., 1973; Pencheff, 1972).

The purpose of this project was to investigate the effects of teaching selected aspects of testwiseness to students enrolled in a reading improvement course. The decision was made to limit the scope of the project to multiple-choice items. The study sought to investigate the following research questions:

1. Do college students demonstrate testwiseness in answering multiple-choice items?
2. Can college students achieve significantly higher scores on a multiple-choice criterion measure of testwiseness following instruction in test-taking skills?
3. What is the effect of test-taking instruction on particular testwiseness strategies?

4. Is there a relationship between reading skills (vocabulary, comprehension, and reading rate) and testwiseness?

Method

Sample

Subjects used in this study were 58 students enrolled in a noncredit college reading improvement course during the Spring semester, 1977. They ranged from freshmen through graduate levels. Most of the students were majoring in the social sciences and humanities.

A pretest-posttest control group procedure was used for the study. Because the investigators were limited to using intact classes, classes, rather than individuals, were assigned to treatment groups. The investigators felt justified in using this procedure because (1) there was no reason to believe that classes differed significantly in composition or ability, and (2) the *Nelson-Denny Reading Test* scores would be used as a covariate to equate for differences in reading ability, if they existed. Both the control group and experimental group consisted of 29 students each.

Procedure

A self-instructional booklet was developed for teaching the selected aspects of testwiseness. The following six strategies were selected for inclusion because of their frequent mention in the literature: (1) eliminate incorrect options; (2) look for repetition or redundancy between stem and options; (3) eliminate similar options; (4) eliminate options which include specific determiners; (5) look for longer, more precise options; and (6) be alert to grammatical inconsistencies.

The general pattern that was used in writing the booklet was to introduce and explain each strategy using an illustrative item and then provide two additional items, with correct feedback, for practice in applying the strategy. After working through all six strategies, a six item review test was taken for added practice in applying the strategies.

The investigators developed a test that could be used to assess students' entering level of testwiseness and that could also be used to assess the effects of instruction in testwiseness. Essentially, the test was to contain items with construction flaws that could be answered more readily by students who were testwise.

A small core of items selected from Millman and Pauk (1969), Ford (1973), and Slakter (1970) was used as a base. The investigators wrote additional items forming an item pool of 60 items, consisting of ten items for each strategy. A panel of judges reviewed the items and identified the five which best represented each strategy. The end product was a 30 item test with five items for each strategy. The items were randomly ordered for the pretest so that all the items pertaining to a particular strategy did not appear sequentially. The same 30 items were randomly ordered a second time to form the posttest.

Results

The results of the data analyses have been organized around the four major research questions which guided this study.

Research Question 1

As an indicator of demonstrated testwiseness, multiple t-tests were performed between the Total Score Mean and the mean Total Chance Score and between the individual subscale means and corresponding subscale chance scores on the pretest. The results of the multiple *t*-tests indicated that all subscale means and Total Score mean differed significantly from the chance scores with the exception of the subscale for Grammar.

Research Question 2

An analysis of convariance was conducted, using the *Nelson-Denny* as the covariate, to test the difference between total posttest means. The results indicated a significant difference between the posttest means for the experimental and control groups, $F(2,55)=54.79$, $p<.01$. The experimental group results were: pretest $M = 16.586$ ($SD = 3.33$) and posttest $M = 24.896$ ($SD = 2.61$). The control group results were: pretest $M = 17.138$ ($SD = 5.13$) and posttest $M = 19.310$ ($SD = 4.34$). The combined results were: pretest $M = 16.862$ ($SD = 4.31$) and posttest $M = 22.104$ ($SD = 4.53$).

Research Question 3

Table 1 presents an experimental and control group comparison of mean percentage of students responding correctly, and gains or losses for individual subscales. The experimental group demonstrated a significant increase over the control group in all six testwiseness strategies.

Table 1

Mean Percentage of Students Responding Correctly to Individual Testwiseness Subscales

	Pre		Post		Gain or	Loss
	Exp	C	Exp	C	Exp	C
Elimination	66.0	75.0	77.0	77.4	11.0*	2.4
Repetition	59.0	66.0	82.0	69.0	23.0**	3.0
Similar Options	63.0	64.0	87.0	70.0	24.0**	5.4*
Specific Determiners	54.0	62.0	93.0	67.0	39.0**	5.2
Length	54.0	52.0	86.0	54.0	32.0**	2.0
Grammar	35.0	39.0	72.0	49.0	37.0**	9.6*

* $p<.05$
** $p<.01$

Research Question 4

The means, standard deviations, and intercorrelations for the four *Nelson-Denny* reading scores and the testwiseness pretest scores for combined groups were computed. With the exception of Vocabulary, there was no significant relationship between reading skills and testwiseness.

Discussion

College students, at least as represented by students used in this study, do possess a significant degree of testwiseness in responding to multiple-choice items. It appears that students were most effective in applying the Elimination strategy, followed by Similar Options, Repetition, Specific Determiners, Length, and Grammar.

Apparently, even though college students possess some degree of testwiseness, -they can significantly improve their ability to apply testwiseness through instruction in test-taking. There was a significant gain for each testwiseness strategy. The greatest gain was found for Specific Determiners with the strategy of Grammar close behind. Elimination produced the least gain, although significant at the .05 level. The investigators believe that instruction had least effect on Elimination because, unlike the other strategies, this strategy involves some knowledge of content. That is, to apply the process of elimination, a test-taker must be able to infer which options are absurd or incorrect. The instructional booklet did not teach the content of the items. On the other hand, the strategies of Specific Determiners and Grammar are most susceptible to superficial characteristics that are independent of content.

There appears to be little relationship between reading ability and testwiseness. The significant relationship between Reading Vocabulary and testwiseness may be due to an underlying relationship between verbal intelligence and testwiseness. Of course, this study cannot clearly confirm this hypothesis, but it would be worthy of further investigation.

References

Callenbach, C. A. The effects of instruction and practice in content-independent test-taking techniques upon the standardized reading test scores of selected second-grade students. *Journal of Educational Measurement*, 1973, *10*, 25-30.

Eakins, D. J., Green, D., & Bushell, D. Effects of an instructional test-taking unit on achievement test scores. *Journal of Educational Research*, 1976, *70*, 67-71.

Ford, V. A. *Everything you wanted to know about testwiseness.* Princeton, N. J.: Educational Testing Service, August, 1973. (ERIC Document Reproduction Service No. ED 093 912)

Gaines, W. G. & Jongsma, E. *The effect of training in test-taking skills on the achievement scores of fifth grade pupils.* Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, 1974.

Gibb, G. *Testwiseness as secondary cue response.* Unpublished doctoral dissertation, Stanford University, 1964.

Jongsma, E. *The effects of instruction in test-taking skills upon student performance on standardized achievement tests* (Final Report of LSU Foundation Research Project). Unpublished manuscript, University of New Orleans, 1975. (ERIC Document Reproduction Service No. ED 114 408)

Kahn, G. C. The effects of testwiseness training on reading scores and on test anxiety. In W. Miller & G. McNinch (Eds.), *Reflections and Investigations on Reading,* Twenty-fifth Yearbook of the National Reading Conference Clemson, South Carolina: National Reading Conference, 1976.

Keysor, R. E. & Williams, D. D. The effect of testwiseness on professional school screening. test scores. Paper presented at the annual meeting of the National Council on Measurement in Education, New York, 1977.

Langer, G., Wark, D., & Johnson, S. Testwiseness in objective tests. In P. L. Nacke (Ed.), *Diversity in mature reading: theory and research* (Vol. 1), Twenty-second Yearbook of the National Reading Conference. Boone, North Carolina: National Reading Conference, 1973.

Millman, J. & Pauk, W. *How to take tests.* New York: McGraw-Hill, 1969.

Moore, J. C. Testwiseness and analogy test performance. *Measurement and Evaluation in Guidance*, 1971, *3*, 198-202.

Moore, J. C., Schutz, R. E., & Baker, R. L. The application of a self-instructional technique to develop a test-taking strategy. *American Educational Research Journal*, 1966, *3*, 13-17.

Oakland, T. The effects of testwiseness materials on standardized test performance of preschool disadvantaged children. *Journal of School Psychology*, 1972, *10*, 355-360.

Pencheff, S. J. *Teaching seven secondary cues of testwiseness with programmed material.*

Unpublished doctoral dissertation, Georgia State University, 1972.

Sause, E. & Grieco, F. *Test sophistication and the inner-city child.* Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, 1973.

Slakter, M. Generality of risk taking on objective examinations. *Educational and Psychological Measurement,* 1969, *29,* 115-128.

Slakter, M., Koehler, R. & Hampton, S. Grade level, sex, and selected aspects of testwiseness. *Journal of Educational Measurement,* 1970, *7,* 119-122.

Tinney, R. E. *The effect of training in test-taking skills on the reading test scores of fifth grade children of high and low socioeconomic levels.* Unpublished doctoral dissertation, University of Minnesota, 1968.

Wahlstrom, M. & Boersma, P. The influence of testwiseness upon achievement. *Educational and Psychological Measurement,* 1968, *28,* 413-420.

Woodley, K. K. *Investigation of testwiseness among C.L.U. candidates.* Bryn Mawr, Pa.: American College, 1972. (ERIC Document Reproduction Service No. ED 062 606)

Yearby, M. *The effect of instruction in test-taking skills on the standardized reading test scores of white and black third-grade children of high and low socioeconomic status.* Unpublished doctoral dissertation, Indiana University, 1974.

VINCENT P. ORLANDO
The Pennsylvania State University

KRISTINA G. HAYWARD
Yankton College

A comparison of the effectiveness of three study techniques for college students

The recent increase of reading-study skills programs at the college level has led to a reexamination of the study skills literature in an attempt to ascertain the effectiveness of various study techniques. Gates (1917) demonstrated that recall of prose and non-prose material was increased when, after reading, subjects attempted to recite (review mentally) everything read. Further investigation into the effectiveness of the recitation-study technique met with mixed results. Peterson (1944) failed to find significant difference between recitation and read-reread groups. More recently, however, Del Giorno, Jenkins and Bausell (1974) adapted the technique to a 455 word prose selection with recitation occurring after each paragraph. They found that in both immediate and delayed recall, subjects who recited after reading recalled significantly more of the passage than those students who read and then reread the passage.

Another attempt to investigate the effectiveness of study techniques has involved assessment of notetaking. A comparative study of the effectiveness of underlining and notetaking while reading conducted by Hoon (1974) resulted in no significant differences in immediate or delayed recall. However, Kulhavy, Dyer and Silver (1975) found superior recall with students who employed notetaking while reading as opposed to students who underlined the text or studied the text. Researchers DiVesta and Gray (1972), and Fisher and Harris (1974) have proposed that notetaking provides an opportunity for the learner to interact with the material which facilitates encoding. However there may be no need for a student to overtly record his/her own ideas for this interaction to occur. Perhaps by providing the opportunity for students to mentally review the contents of the passage, as in the Gates recitation technique, the same encoding function would be served.

The purpose of this study was to assess the effectiveness of three study techniques as measured by a recall test of an 800 word passage on both immediate and delayed tests. The three study techniques were: 1) reading and then rereading the passage; 2) reading the passage with periodic mental review; and 3) reading the passage with notetaking.

Method

Subjects

The subjects were university juniors and seniors enrolled in an elementary education methods course at The Pennsylvania State University who volunteered to participate in the study. Forty-eight students were randomly assigned to one of the three treatment groups.

Materials

The stimulus material used in each of the treatment groups was an 800 word passage about vitamins developed by Sagaria and DiVesta (1977). The passage was divided into ten paragraphs of seven sentences with each sentence containing one idea. A recall item was constructed for each sentence. The resulting 70 questions thus represented all the factual concepts presented in the passage. Through the use of a computer processing program, 40 questions were randomly selected from the domain of the 70 items to make up a separate test for each subject. Each set of 40 items equally sampled the information from the ten paragraphs. Once selected, the order of the questions in each recall test was randomized to reduce the likelihood that an item could be answered because it duplicated the sequence of the material within the passage.

Procedure

The data for the study were collected during two sessions. During each session students met in small groups of four to eight. At the beginning of the first session the stimulus material was disseminated and students were advised that upon completion of the study task they would be administered a 40 item recall test. They were also informed that in one week's time they would return for a similar test on the same material.

Oral instructions on the study technique to be used were given to each treatment group. Students in the read-reread group were instructed to read through the material as if they were reading an article, not stopping to mull over ideas. Upon completing the passage they were instructed to reread it in the same manner. Students in the read-mental review group were instructed to read the passage one paragraph at a time. After reading each paragraph, they were to look away from the passage, attempt to recall the information in the paragraph from memory, and then look back at the paragraph to check to see if they had missed any important information. They were instructed to continue this process until the paragraph was mastered. Students in the read-notetake group were instructed to follow the same basic procedure as the read-mental review group. However, they were to write the information recalled after each paragraph.

All students were required to record the time they used to read/study the passage from a digital clock at the front of the room. Upon finishing the reading/studying, each student was given his/her 40 item short answer test. During the second session, one week later, the students were administered a different set of 40 items randomly selected from the domain of 70 items.

Analysis

In order to assess the performance of the three study groups on the immediate and delayed recall tests, a 3 X 2 analysis of variance (study-condition by test-condition) was computed. In addition, the study time factor (minutes used to study) was also analyzed with a one-way analysis of variance with the three treatment groups serving as a between subjects factor. Level of significance was set at the .01 level, and follow-up pairwise comparisons were made via the WSD procedure.

Results

The study condition by test-condition analysis of variance revealed significant main effects for both study condition, $F(2,45) = 6.44, p < .01$, and test condition, $F(1,45) = 143.17, p < .01$, and a significant interaction $F(2,45) = 8.46, p < .01$. In the immediate test condition the follow-up comparison of study condition effects using the WSD procedure indicated that there was no

253

significant difference between the performance of the read-notetake and read-mental review groups (mean number of correct responses = 29.9 and 29.0). However both were significantly different from the read-reread group (21.0). In the delayed test condition, although the same basic trend in means occurred among the read-notetake, read-mental review, and read-reread groups (20.0, 20.6 and 17.6) there were no significant differences among any of the groups.

The one-way analysis of variance of the study time factor revealed an overall significant main effect, $F(2,45) = 61.08$, $p < .01$. The follow-up comparison using the WSD procedure indicated that the read-notetake, read-mental review, and the read-reread groups (mean study time in minutes = 26.92, 17.38 and 8.75 respectively) were each significantly different from one another.

Discussion

The results indicated that on an immediate recall test, the study techniques of notetaking and mentally reviewing produced significantly superior results to the rereading technique. In the delayed test condition, although the trend in means was the same, there were no significant differences. Significant study time differences existed among all 3 groups with notetakers using the most time and rereaders the least.

For immediate recall it appears that both notetaking and mentally reviewing the passage are effective means of studying. However, it is inappropriate to ascertain effectiveness without attention to study time. If it were assumed that difference in study time could account for difference in the amount of information recalled then a direct relationship between the amount of information recalled and the amount of study time should exist. However, this relationship is not present. Although the notetakers spent significantly more time studying, their immediate recall was not significantly different than those students who mentally reviewed the contents of the passage. We can conclude, therefore, that for immediate recall the mental-review study technique was the most efficient study method. Further, it appears that the encoding function served by notetaking is also served by mentally reviewing each paragraph read.

These effects do not persist in the delayed recall condition. Neither the notetakers nor the mental reviewers performed significantly better than those students who merely reread the material on the delayed test. Another function served by notetaking, according to some researchers (DiVesta & Gray, 1972), is that of providing a means of external storage of essential information from the study passage. In this experiment subjects were not allowed to make use of this opportunity. By providing students review time under the delayed condition a more typical study situation would be replicated and hence more definitive conclusions regarding study effectiveness could be made.

In summary, the read-mental review study technique was the most effective technique for immediate recall with the consideration of the study time factor. Conclusive evidence regarding study technique effectiveness for delayed recall was not provided by this study but was recommended as an area for further research.

References

Del Giorno, W., Jenkins, J. & Bausell, B. Effects of recitation on the acquisition of prose. *Journal of Educational Psychology*, 1974, *67*, 293-294.

DiVesta, F. J. & Gray, S. Listening and notetaking. *Journal of Educational Psychology*, 1972, *63*, 8-14.

Fisher, J. & Harris, M. Effects of notetaking and review on recall. *Journal of Educational Psychology*, 1973, *65*, 321-325.

Gates, A. Recitation as a factor in memorizing. *Archives of Psychology*, 1917, *40*, 22-64.

Hoon, P. Efficiency of three common study methods. *Psychological Reports*, 1974, *35*, 1057-1058.

Kulhavy, R. W., Dyer, J. W. & Silver L. The effects of notetaking and test expectancy on the learning of text material. *Journal of Educational Research*, 1975, *68*, 363-365.

Peterson, H. A. Recitation or recall as a factor in the learning of long prose selections. *Journal of Educational Psychology*, 1944, *44*, 220-228.

Sagaria, S. D. & DiVesta, F. J. Learner expectation induced by adjunct questions and the retrieval of intentional and incidental information. *Journal of Eduicational Psychology*, in press.

DR. HELEN C. COVINGTON
North Harris County College

DR. LEE MOUNTAIN
University of Houston

The competencies that junior college chairpersons expect from their reading course graduates

Remediation courses in reading and study skills are rapidly expanding within the junior college curriculum. These courses have been instituted to stem high student attrition by preparing students to succeed in either regular college courses or in their areas of vocational interest (Ahrendt, 1975; Sweiger, 1972; Roueche, 1967).

The diversity in organizational structure of these remedial programs is evident from a review of the literature in the field (Covington, 1977). Some colleges have just one reading course while others have as many as four or more courses. Research in the field of developmental studies (Gwynne, 1973; Sweiger, 1972) has shown the predominant goal of these courses to be one of preparing the student for "his years ahead in college" (Sweiger, p. 7).

With such a diversity of courses that aim for student success in regular college courses or "immediate personal, vocational competence" (Huffman, 1975, p. 4306a), the need arose to ask of reading department chairpersons, "What, precisely, are the competencies you expect from your reading course graduates?" The necessity for such a question is apparent. Is there a core of reading skills that all reading graduates should possess for basic literacy, personal satisfaction, and minimal vocational competence? Beyond this core, are there other reading skills in which reading course graduates should be proficient for success in regular college courses?

A study was conducted to determine what chairpersons considered to be (1) the major components of junior college developmental reading courses, (2) competencies expected from reading course graduates, and (3) the areas in which such courses could be improved. An additional analysis was conducted to determine whether there were significant differences in the responses given by chairpersons responsible for one of two remedial reading courses and chairpersons responsible for three or more such courses.

Method

Questionnaires were mailed to fifty-nine chairpersons of junior college reading and study skills programs in Texas. Twenty-eight of these chairpersons responded. Each was asked to rate twenty-one items according to the importance of the item in the developmental reading program. A five-point scale from Highly Important to Unimportant was used for this purpose. Percentage distributions were obtained for each of the five categories.

Ultimately, these categories were reduced to two: the positive category consisting of the responses obtained from Highly Important and Important, and the negative category consisting of responses from Somewhat Important, Relatively Unimportant, and Unimportant. Although the items cited as Somewhat Important could have been positively categorized, the researchers chose to include in that category only those items clearly delineated as *strongly* positive ones by chairpersons.

The binomial distribution (Parsons, 1974) was used to determine critical values, thereby establishing the groupings of course components designated as *Least Important*, *Important*, and *Most Important*, as shown in Table 1.

In addition, chairpersons were asked to write an essay in which they commented on the competencies expected of their reading course graduates and cited the improvements and amendments, if any, that would enhance the present reading program.

Finally, in order to determine whether the responses of chairpersons with smaller departments were significantly different from the responses of chairpersons with larger departments, the twenty-eight chairpersons were divided into two groups. Group A consisted of seventeen chairpersons whose programs consisted on one of two courses. Group B consisted of eleven chairpersons with three or more courses. An analysis of variance was conducted to determine whether there were significant differences in the responses of the two groups to each item.

Results

Six of the twenty-one reading and study-skills items were deemed *Most Important* components of a reading course. Four of these six items were in the area of contextual comprehension. Seven skills were considered to be *Important*. Eight were deemed *Least Important* (see Table 1).

Junior college reading chairpersons expect the graduates of their courses to be able to demonstrate competencies in at least four of the course components designated as *Most Important*, namely the skills of comprehension in context.

The answers to the question, "What improvements and amendments, if any, are needed in the program?" were all provided from the responses in the chairperson's essay. In the area of program philosophy, chairpersons recommended an overall improvement in communication between staff and students so that all understand the importance of reading in everyday life, in subject area courses, and in career selection. Further, chairpersons suggested that communication between staff members and remedial students would be enhanced if the staff members would more frequently approach remedial students with confidence as people who can become proficient in reading.

For improvement in student self-image, chairpersons felt that spiraling courses would aid the remedial student better than separate departmental remedial courses because disciplinary reinforcement and counseling could be more easily integrated. The courses in the program, they suggested, should be arranged from basic reading upward, with special multi-disciplinary labs and learning centers staffed by a team of professionals, para-professionals and tutors.

Other aspects of the program should include the granting of credit, pass-fail grading, small teacher-student ratios, non-registration and continuous enrollment for lab students, and flexibility within the program structure.

The teaching techniques suggested would include diagnostic-prescriptive methods, continual assessment, independent reading time within the program, modular lesson planning, group as well as individual (self-paced) instruction, and more creative teaching.

Other improvements suggested were for more diversified, multi-level, self-paced, programmed, and ethnic materials. A final suggestion was that computerized help in selecting materials appropriate for certain deficiencies be made available to reading and study skills personnel.

There were no significant differences noted in the type of response given by chairpersons of large departments and chairpersons of small departments.

Table 1

Level of Importance of Reading Course Components

		LEVEL OF IMPORTANCE		
CATEGORIES	COMPONENT	MOST IMPORTANT	IMPORTANT	LEAST IMPORTANT
CATEGORY I STUDY SKILLS				
A. Listening and Notetaking	1. Notetaking			X
	2. Outlining			X
	3. Listening Skills	X		
B. Research	1. Dictionary		X	
	2. Library Skills			X
	3. Research Skills for Content Areas			X
C. Test Taking	1. Test-Taking techniques		X	
	2. Test-Questions Analysis			X
D. Textbook Study	1. Textbook Organization		X	
	2. Study Methods (SQ3R, etc.)	X		
CATEGORY II READING SKILLS				
A. Comprehension in Context	1. Basic Comprehension, Main Idea, Details, etc.	X		
	2. Following Directions	X		
	3. Advanced Comprehension			X
	4. Vocabulary in Context	X		
	5. Syntactic Understandings	X		
B. Conceptual Thinking	1. Induction			X
	2. Deduction		X	
	3. Concepts (Cause & Effect, etc.)		X	
C. Speed in Reading	1. Flexible Reading Rate		X	
	2. Speed Reading			X
D. Phonics	1. Phonics		X	

Conclusions

Significantly, the skills deemed *Most Important* by chairpersons illustrate their feelings that the major course components (and ultimately, the competencies of the remedial graduate) are, in the main, the mastery of basic reading skills. The fact that chairpersons of small as well as large reading departments gave responses that were not significantly different further illustrates that certain basic skills are seen across the board as the major components of remedial reading courses.

Some of the more advanced skills in the areas of "Research" and "Conceptual Thinking" were cited by chairpersons as *Least Important* for a remedial program. The "Conceptual Thinking" skills and the entire list of Study Skills, however, could be of extreme importance in aiding students for future college success. Half of the Study Skills items were rated *Least Important*, with only two of these skills attaining top priority in the reading program. Until lecturing is replaced by other methods of disseminating content information, chairpersons may need to teach notetaking and outlining. They can thereby prepare remedial students to take class notes and rewrite them in orderly outlines for better intellectual processing.

If the findings of this study truly reflect the expectations that junior college reading chairpersons hold regarding the competencies of their reading course graduates, we suggest that they consider moving toward broader expectations.

References

Ahrendt, K. M. *Community college reading programs.* Newark, DE: International Reading Association, 1975.

Covington, H. C. *Analysis of three sets of expectations concerning reading and study skills programs in selected Texas junior colleges.* Unpublished doctoral dissertation, University of Houston, 1977.

Gwynne, M., Canine, J., & Quigley, M. *Evaluation of remedial programs in community junior colleges.* New York: Teachers College, Columbia University, 1973.

Huffman, J. A. A study of perceived and desired aims and objectives of developmental programs in Illinois Community Colleges (Doctoral dissertation, St. Louis University, 1975). *Dissertation Abstracts International,* 1976, *36,* 4305A-4306A. (University Microfilms No. 75-858)

Parsons, R. *Statistics for decision makers.* New York: Harper and Row, 1974.

Roueche, J. E. *Experimental programs in the junior college.* Los Angeles, CA: 1967. (ERIC Document Reproduction Service No. ED 013073)

Sweiger, J. Teaching reading as a reality at the community college. *Twenty-Second Yearbook of the National Reading Conference,* 1973, 111-115.

EDWARD G. SUMMERS
ANNE FORESTER
SHARON JEROSKI
University of British Columbia

1975-1976 review of research in college-adult reading

The last review of research in college-adult reading was published in 1975 and covered the area through the end of May 1974 (Blanton and Smith, 1975). This review includes research which appeared in the literature for the two years 1975 and 1976 and covers approximately June 1974 through May 1976. The reports reviewed cover the published literature directly concerned with research in college-adult reading except in some instances where allied references have been included in explanatory sections or discussions of issues which span more than a short time period.

The articles reviewed have been organized within the following broad conceptual framework.

Surveys of College-Adult Reading Programs
Characteristics and Needs of College-Adult Readers
Evaluation of Programs and Instructional Methodology
Rate and Flexibility
Readability, Materials/Textbooks
Test Development, Evaluation and Use
Word Recognition
Sentence Comprehension
Learning From Prose

Surveys of College-Adult Reading Programs

Smith, Enright and Devirian (1975) reported results from a nationwide survey of learning and study skills programs using an all-inclusive definition of learning center. Results indicated that 61% of respondents were running learning study skills programs with an additional 9.3% indicating plans to develop a program.No established pattern of administrative responsibility has yet emerged and course credit is still a vital issue. Maxwell (1975) presented a case study sketching the development, program philosophy and operation of a university learning center.

Reading and study skills programs in selected Rocky Mountain colleges and universities were surveyed by Warren (1975) who found 84% of institutions operating programs with 9% planning courses. Two year institutions offered fewer courses than did four year ones but served more students. Over three-fourths of the courses were offered for credit with wide variation existing in use of machines, materials and instructional procedures. Few instructors had formal academic training in reading. Programs were aimed more toward minority and academically troubled students than toward total school populations.

Manikas (1974) analyzed past and present remedial reading programs for underprepared black junior college students, identified components from 30 successful programs, and presented a model remedial reading program for such students.

Survey data were analyzed by Haviland (1974) to determine the status of publicly and privately administered illiteracy programs in Great Britain. Programs have grown rapidly, particularly since 1967, and awareness of the problem is increasing. However, current provisions for dealing with the complexity and magnitude of the problem are inadequate.

Kinnamon (1975) reported results of a questionnaire survey sent to 160 commercial TV stations in the U.S. and Canada to determine type and number of programs offered which are designed specifically to teach adults to read. Program characteristics are described for the few stations offering them.

Appreciation is expressed to Donna Dartnell and Catherine Tolsma for aid in collecting and organizing data and to Catherine Tolsma for manuscript typing.

Acknowledgement for funding and released time is made to the UBC Faculty of Education and the Reading Education Department.

260

College Readers – General Characteristics

In studying long-term trends in the reading ability of college freshmen Larsen, Tillman and Cranney (1976) analyzed the results of a decade of testing for all incoming freshmen at the University of Florida. As measured by the DRT, reading ability was highly stable for the ten year period 1960-1970; even a sizeable increase in admissions requirements (achievement tests) in 1962 did not result in statistically significant changes in reading score trends.

Ambrosino, Brading and Noval (1974) tested the assumption that entering medical students have successfully developed adequate reading skills using NDRT scores for 268 students entering Albany Medical College as the classes of 1975-1976. Analysis of scores and careful review of admissions records revealed that in at least half the cases of academic failure reading deficiencies played a prominent role. Denne (1975) also examined problems of admission and academic success of medical students using McGraw-Hill Basic Skills System Reading Test data for 85 students at the University of Utah. No correlation was found between reading test scores and achievement in the first two quarters of work in medical school, and the decision was made not to include a reading test as part of the admissions procedure.

Hornsby (1974) investigated differential language behavior between vocational and academic students using the Oral CIA Language Inventory (free reporting), the DRT, the Otis, Purdue Pegboard and General Aptitude Test Battery for 60 randomly selected subjects enrolled in programs in a Mississippi junior college. ANCOVA, covarying on I.Q., indicated no significant differences in paradigmatic responses and significant differences in favor of the academic group in reading ability. Correlational analysis yielded nonsignificant correlations between paradigmatic responding, reading, I.Q. and performance measures.

Case study methodology was employed by Wiggins (1975) to gather data on negative environmental factors in home and school, intellectual level, interests, and personal and social adjustment for Black male remedial reading freshmen in a small Southwest church college. It was concluded that disabilities and specific deficits represent a cumulative educational deficit.

Ford (1974) examined the relationship between SES, mode of linguistic expression, attitude toward and interest in reading, level of aspiration, community environment, age, sex, and I.Q. with reading ability utilizing correlational and multiple regression techniques based on scores of 60 first year Ontario college students. I.Q. made the greatest contribution to reading ability and the two major predictors of college achievement were characteristics of the home and reading ability. SES variables, attitude, I.Q., and some of the interest variables were significantly related to reading skills. Poorer readers speech represented a restricted code while better readers exemplified the elaborated code.

The variables of life change measured by the Social and Collegiate Readjustment Rating Questionnaire, NDRT achievement, course-hour work load, and selected attitude factors were analyzed by Henard (1976) as predictors of year end academic performance for freshmen at North West Texas Community College. Descriptive analyses and ANOVA indicated reading achievement was an effective predictor of GPA, course load and selected attitude factors while life change and the interaction of reading ability and life change effectively predicted only course-hour load.

Corlett (1974) correlated scores on the Library Orientation Tests, the SSHA and sex with GPA using 81 freshmen English students at a small metropolitan university. Intercorrelations among all possible permutations of the predictor variables and the criterion measures suggested the tests were relatively independent of each other. The Library Orientation Test appeared to be valid for forecasting success in college. Relatively little variance was added with the use of the SSHA; women tended to achieve at a higher level than men and also achieved a higher predicted GPA.

Solis (1974), using 500 junior college students, examined the effects of ten independent variables on residual reading gain using stepwise multiple regression, t tests and chi square analysis and found unfulfilled reading potential, number of semesters in reading instruction and initial reading ability to be relevant predictors.

GPA. Discussions of the use of GPA as the criterion variable in characterizing needs and assessing effects of reading instruction continue to be reported in the literature (Farrell, 1975). Wilson (1974) used a stepwise multiple regression program to determine the statistical relationship between college GPA and reading index of freshmen students at Florida junior colleges. Findings revealed a positive and significant relationship between reading ability and GPA. Pedrini and Pedrini (1975a, b) evaluated the contradictory research examining the relationship between reading ability and GPA. In research using 120 college students (NDRT and course grades), it was found that 35% of the grade variance was accounted for by vocabulary scores. Course success was related to reading achievement only insofar as reading was emphasized as being important in the course. Haselton (1974) used ANOVA and correlational analysis to examine the relationship of specific reading and study skills, and faculty perception of those skills, to GPA in each of 10 schools at Oregon State University. A distinctly different pattern of reading and study skills emerged for each of the ten schools. Paragraph comprehension correlated significantly with GPA in 9 out of 10 schools and scores for skimming and scanning correlated more positively with GPA than any other scores although identified by the majority of instructors as least important.

Adult Readers – General Characteristics

DePietro (1975) attempted to predict achievement, attendance and dropout incidence in adult basic education projects for 397 students in four programs randomly selected from 16 Right to Read projects in Northeastern urban areas. The predictor variables of enrollment, age, sex, and formal educational attainment did not predict attendance or attrition; teacher directiveness did predict literacy achievement.

The oral reading performance of samples of functionally illiterate adults and second and third grade developmental readers were compared by Russell (1974). Point biserial correlation coefficients revealed significant similarity of error patterns between the two groups for all categories except syllabication when reading at the frustration level.

Mizer (1975) compared university students 50 years of age or older with an analagous group not enrolled in formal education. MANOVA indicated significant differences in life satisfaction, reading proficiency, and interpersonal values but not for selected background characteristics.

Based on a follow-up study of 40 adult males who were clinically diagnosed as dyslexic in childhood, Frauenheim (1975) reported that results of oral, comprehension and vocabulary reading tests and arithmetic and spelling achievement measures, administered an average of 10 years 3 months after diagnosis, revealed that subjects had remained severely retarded in reading (mean reading grade score 3.6), with concomitant effects on all aspects of life adjustment.

Personality Characteristics

General. Trachtman (1975) tested 50% of the incoming freshmen class of disadvantaged minority students entering a special program to determine the relationship of cognitive, demographic and certain intellectual achievement variables to academic performance. Step-wise multiple regression analysis revealed reading, attitude toward authority, dogmatism and selected items of internality-externality to constitute the best combination of predictors, accounting for 35% of the variance.

Tillman, Millot and Larsen (1974) analyzed the results of five studies which explored the relationship between personality type and reading test scores, hours worked at the reading center, enrollment in voluntary reading programs, GPA of marginal students and reading test scores. Two tentative conclusions were reached: persons who state a preference for dealing with abstractions and symbols tend to be more effective readers; those who prefer to please other people tend to benefit more from individualized helping situations.

The interrelationship of study skills and personality variables was studied by Rutkowski and Domino (1975) using data from 201 university freshmen who took the California Psychological Inventory, the SSHA and the Scholastic Aptitude Test. Results of the factor analysis cautiously support the independence of study skills, personality variables and aptitude scores.

Hoiberg, Hysham and Berry (1974) replicated a previous study and compared discharge information on Navy recruits assigned to a remedial training program with that for a control group to assess the psychiatric implications of illiteracy. The percentage of neuropsychiatric and disciplinary discharges during the first year of active duty was significantly greater for the illiterate group.

Treppa (1973), using the Interpersonal Check List, terms students who enroll in college study skills courses to be more self-rejecting, passive, and to perceive their parents more negatively than a control group of subjects.

Self-Concept. Musik (1974) analyzed pre and post treatment data for the DRT and a researcher developed self-concept instrument for community college students randomly selected from a reading improvement course and assigned to one of three treatment groups differing in their relative emphasis on reading or self-concept training. They concluded that the combination of reading and some training in self-concept was most promising for reading improvement.

Lund and Ivanoff (1974) utilized 227 students to determine the difference in measured self-concept among college freshmen blocked by reading ability, sex and enrollment in a reading skills program. All variables, particularly course enrollment, were sensitive to self-concept variation.

Internal-External Locus of Control. Allen and Harshbarger (1977) compared the performance of internally and externally oriented college reading improvement students taught in a student-controlled or an instructor-controlled program. ANOVA and pairwise comparisons revealed that internally-oriented students in classes highly controlled by the instructors made significantly smaller reading achievement (DRT) gains than did all other students. Drummond, Smith and Pinette (1975) investigated the extent to which I-E construct related to student achievement in an individualized community college reading program. Results for 30 male students revealed greater achievement for externally oriented students.

Interests

McCreath (1975) assessed reading habits and interests and their relationship to reading improvement for 89 disadvantaged junior college students enrolled in reading improvement courses. Generally, students were not regularly involved in reading activities. Newspapers and magazines appeared to be fairly regular reading with pleasure reading, discussion of reading and use of libraries scoring low. Patterns and habits were independent of age.

Several studies assessed various factors related to newspaper readership and readability among adult populations. Penrose et al. (1974) used a sample of 1,130 randomly selected respondents in North Carolina to compare the profiles of newspaper and non-newspaper readers with the results of a study conducted in Wisconsin a decade earlier. Results suggest the same type of people (low SES, rural, low education) continue to be the non-newspaper readers. Schweitzer (1976) analyzed young people in a metropolitan area to determine the factors involved in subscription or non-subscription to two daily newspapers. Chi square tests revealed statistically significant differences between subscribers and non-subscribers on marital status and occupation; lifestyle, mobility and educational variables were not related. Bryant, Currier and Morrison (1976) analyzed a sample of 900 adults from the Toronto metropolitan area to determine if lifestyle rather than demographic factors relate to newspaper choice, resulting in a new technique for defining readership.

Attitudes

Stevens (1975) assessed change scores on a reading test and a semantic differential instrument for adult students in eight reading classes using a correlation matrix. Significant relationships between initial attitudes toward reading and reading skills improvement, but not between attitude change and reading score changes, were noted.

Physiology

Reichurdt (1977) used a polygraph to measure the physiological reactions of reading disabled students (median age 21, 1st to 30th percentile–Davis Reading Test) and reported statistically significant differences between their reaction to oral and silent reading tasks. Gross abnormalities in respiratory patterns were found to exist among long-term reading handicapped students. From the standpoint of arousal, two distinct groupings emerged: the lowest scoring readers were found to be hypo-tensive while those scoring somewhat higher were hyper-tensive.

Wark (1974) compared validly test-anxious with validly test-cool psychology students in terms of their heart rate patterns while resting and while reading to answer test questions and reported a clear distinction on heart rate pattern between the two groups.

Evaluation of Programs and Instructional Methodology

A lively interest continues in the evaluation of program effects and instructional methodology in college-adult reading, particularly among doctoral candidates. Two-thirds of the research reviewed in this section was reported in doctoral dissertations.

Romine (1974) polled students and faculty in 29 junion-community colleges in 15 states using a 70-item instrument designed to assess attributes of an effective instructional climate. Factor analysis produced an instrument consisting of 41 attributes which make a significant contribution to instructional climate. Characteristics are grouped under the headings of instructor preparation, planning and organization, presentation, and assessment and outcomes. Intercorrelations and reliability analysis reorganized the data further to produce nine major personal and academic characteristics which typify an effective college instructional climate.

Robin (1976) reviewed the nature of behavioral instruction; its effectiveness in terms of academic achievement, retention, attitudes, study time, and withdrawals; and assessed the contribution of various components of the individualized method including self-pacing, unit perfection requirements, stress on writing, use of proctors, study objectives, grading, lectures, and student/component interactions.

Reading Programs

Fairbanks (1974, 1975) examined 87 reading study skills programs and developed a system of rating the degree of research control utilized in reports of their effectiveness. Fairbanks concluded that poor research design and control may well have inflated results concerning reported effects of college reading programs on GPA. She suggested that variables such as content and mode of operation of the program, response to program variables, student personality variables, entry reading skills, scholastic aptitude and achievement, academic curriculum and staff enthusiasm be studied further. Parham (1975) studied the relationship of persistence in participating in a college level reading course to six variables including GPA, sex, personality, entrance reading proficiency, general scholastic ability at entrance, and the type of curriculum being pursued by the student. Persistant and non-persistent students did not differ on any of the variables.

Burgess, Cranney and Larsen (1976) investigated short and long term effects on GPA of a free, voluntary noncredit reading improvement program using 46 freshmen as a treatment group and the entire remaining freshman class (N = 3,000) as a control group. Characteristics of the voluntary program were similar to those noted for successful programs in the Fairbanks (1974, 1975) studies. Significant differences in GPA's favored the treatment group at the end of the first and second term.

The effects of three treatments (counseling, academic aid, no treatment) on college grades, reading scores and attrition for 71 low income freshmen were studied by Cochran (1975). Significant positive relations were found between reading scores and GPA; however, there were no differences for either NDRT reading change scores or GPA.

Hoiberg, Hysham and Berry (1974) compared 1,517 enlistees who attended an academic remedial training program with a matched comparison group. Higher values of final reading level, the AFQT, and educational attainment were all predictive of military effectiveness. The remedial training program did aid many men to stay on active duty to perform effectively in the Navy.

Carpenter and Jones (1975) used the NDRT to measure gains in comprehension and rate for students enrolled in reading classes at Clemson University. Significant gains were found for subjects taking reading classes. Large gains in comprehension were attributed to the sequential acquisition of separate comprehension sub-skills and their subsequent integration and application to content materials.

The influence of reading classes on the reduction of phonic disabilities of first year students was studied by Littrell (1976). Experimental and control groups were analyzed consisting of students enrolled in English and elective reading classes in a Kansas junior college. ANCOVA results of pre and post scores on the California Phonics Survey showed significant gains for reading class students as compared to controls enrolled in English classes and significant gains for the reading classes in comprehension as measured by the Diagnostic Reading Test.

Special Admission Students. Papandrea (1975) compared self-concept and achievement change scores of learning laboratory participants who were in the lowest quartile of their graduating classes with a control group of potential learning laboratory students. ANCOVA, correlational analyses and t tests indicated mixed results in comparing the two groups on reading achievement, arithmetic, self-concept and English. No significant relation was found between self-concept and academic achievement. Rakes, Lintz and Payton (1974) found, when comparing student achievement and attitude in applying competency based and traditional teaching programs to four groups of predominently Black students, that competency based classes were superior to traditional classes on the California Reading Test. Student attitude toward competency based teaching was positive but the costs of operating such classes were three times those for traditional classes.

O'Donnell (1974) compared academic skill gains of 42 Black high-risk college freshmen, attending a summer program, assigned to tutoring in the content of the traditional course or tutoring (11 one hour sessions, college upperclassmen) using a syntactically oriented psycholinguistic program. MANOVA and t tests indicated no significant differences between the groups in reading achievement, syntactic maturity of written composition and attitude. However, Whittaker (1974) found a linguistically based instructional method to be significantly more effective than a programmed method in improvement of comprehension but not vocabulary for a 16 week reading and study skills course.

Weiss (1974), using ANCOVA, found differences in favor of a language experience approach over a traditional approach for total reading and comprehension but not vocabulary. Kahn (1975) compared reading achievement of students receiving a joint program of applied psychology and communication skills for two hours three times a week for fifteen weeks with that of students receiving only communication skills instruction for three one hour sessions each week for fifteen weeks. No differences appeared on the posttest analysis although the psychology plus communication students read more than twice the number of books and experienced changes in attitude.

Smith (1975) compared the gains of 75 junior college students enrolled in a diagnostic/prescriptive program with a traditional control group. No significant differences were found between reading gains of the two groups. However, both groups made significant gains during the eleven week instructional period. Rivera (1974) found that students enrolled in an eight week program of individualized tutorial assistance in reading and study skills while failing to attain significantly higher reading scores or lower attrition rates than control students not receiving tutorial assistance, did achieve significant improvement in GPA.

ABE Programs. Knowles (1975) compared demographic, personality and achievement data of students in a state-supported county-wide ABE program with those enrolled in the volunteer tutor Right to Read program and concluded. The two programs served essentially the same highly variable populations. Personal rather than media recruitment was preferred by participants and provision of transportation significantly affected involvement for functionally illiterate adults. Using a sample of 443 subjects, Lamb (1975), working with adults in four Florida counties (ages 16-56, reading level 3-8), found that general instruction increased critical reading ability while maturation in the absence of instruction did not. Hutchison (1974) found that an experimental program with lessons divided into teaching, practice and application was significantly more effective than large group instruction utilizing commercial materials over a fifteen week period on both achievement and attendance measures but not on vocabulary or oral reading.

Spears (1974) found significantly better academic growth and retention for an individualized, programmed laboratory program than for formal adult education programs involving individualized, small and large group instruction in self-contained classrooms in neighborhood schools.

Johnson (1975) concluded that CAI programs were effective in ABE in a study involving seven program components (content, question type, question hierarchy, terminal choice, branching, timing, feedback) and five learner characteristics (history of failure, negative attitude toward school atmosphere, immaturity, dependency, suspiciousness).

Tinkle (1974) found significant advantages for massed over spaced instruction using prison farm inmates, even though both types of instruction were effective with illiterate inmates.

Instructional Methodology

Gnewuch (1974) assessed the effects of a program designed to increase awareness of words and word meaning based on Dale's principles. Change scores for a 12 week program for 318 experimental and 89 control university reading and study skills students on the NDRT indicated that teacher training and experience contributed more than ten times the amount of variance contributed by treatment or classes. Wolfe (1975) found no differences in writing complexity or usage for three vocabulary development programs: teaching a specific reading vocabulary with practice of words in sentences, teaching a specific reading vocabulary with practice of words in multiple choice exercises, and no specific vocabulary instruction. The effects of modified paperback scanning and vocabulary extension instruction on NDRT (vocabulary) and Carver-Darby Chunked Reading Test scores were investigated by Baer (1975). ANOVA, using 75 volunteer undergraduate psychology students, indicated the paperback method was superior for reading accuracy but not for reading vocabulary, comprehension and rate.

Henderson (1975) analyzed pre and posttest scores on reading rate, vocabulary and comprehension, self-concept, reading interest variables, and self-expectation for eight intact community and junior college classes randomly assigned to prescriptive or personalized reading improvement instruction. Stepwise discriminant analysis indicated no significant pre or posttest differences with the exception of males in the prescriptive group over males in the personalized group. Allen (1976) considered three levels of student control (minimum, moderate, maximum) and found no significant differences in reading achievement (DRT), attitude toward reading (Estes) or attitude toward the course by treatment group, and no relationship between student personality orientation (Rotter I-E Scale) and reading achievement in any of the college reading groups. The reading achievement and interest gains of two year college students taught by a personalized technique (cognitive field theory) and a prescriptive technique (behavioristic theory) were investigated by Henderson (1976). No differences favoring either group were found for either dependent measure. Mahoney (1974) randomly assigned 119 university students taking reading and study skills classes to a traditional teacher-centered approach, a didactic tutorial approach, a facilitative tutorial approach, and no instruction. ANOVA, post hoc comparisons and Chi square indicated all experimental groups superior to the control group on NDRT scores but not GPA or mean verbal GPA.

Kistulentz (1976) compared academic course achievement scores, GPA and NDRT change scores of students enrolled in a community college reading improvement course and assigned to traditional instruction or sections where vocabulary and comprehension exercises were derived from composition, mathematics and science texts. Both groups made significant gains in NDRT reading achievement. There were no differences in grade point averages, reading improvement scores and achievement in the writing and mathematics courses; however, there were significant differences in favor of the experimental group in achievement in the science course.

Bain (1974) compared the reading gains of 60 experimental and 60 control students enrolled in a college reading course who simultaneously listened and read SRA materials and read without tapes. NDRT gain scores using t tests indicated no significant difference by treatment. The reading comprehension gains of students stratified by reading achievement and assigned to reading, listening, and combined treatments, were analyzed by Thomas (1975). ANOVA results for the original experiment and three replications indicated significant main and interaction effects. The dual mode was more effective than unimodal presentation and listening more effective than reading. The interaction suggested that the lower the initial reading level, the more effective the dual mode.

Gurrola (1975) applied ANOVA to gain scores on teacher constructed tests for treatment groups taught study methods based on different combinations of SQ3R components and found a treatment effect for study efficiency and a reading level effect for study effectiveness with a population of university freshmen.

The effects of proportional amounts of time spent in diagnosis and in instruction on achievement gain for students were examined by Livingston (1975). Increased diagnosis facilitated achievement gains for average and below average but not for above average groups following ten weeks of instruction.

College-Adult Reading Personnel

Phlegar (1975) surveyed the attitudes of faculty and examined effects of an inservice reading program for faculty in a community college. Results for all faculty members, and responses for selected students and faculty members, suggested: inservice affected teacher attitudes toward content area reading but did not change student attitude; female teachers were more conscious of the importance of reading and more experienced teachers more aware of the complexity of reading; attitudes of teachers and students were similar and few differences appeared between academic and occupational-technical teachers; married students were more serious about reading and their classes; and females enjoyed reading more than males, although males reported spending more free time reading about the subject.

Schnell (1974) provided a profile of the community college reading teacher based on questionnaire returns from 47 members of the IRA special interest group for two year colleges. The composite community college reading teacher was described as being female, forty-five or older, highly educated and experienced in teaching. She feels her college training was only adequate and that experience has been more valuable than course work. Ramig and Dulin (1975) described a one-year reading methods seminar offered by the University of Wisconsin-Madison for adult basic education teachers including program goals, administrative considerations, format of the course and evaluation in which students rated the course and rank ordered, in terms of usefulness, ten areas of course emphasis.

Rate and Flexibility

Perception

In examining factors which contribute to the difference in reading speed among young and older age readers Marcel (1974) ruled out visual processing factors such as interfixation eye movements, number of regressions, fixation time and larger eye-voice span (EVS) and hypothesized that a prime cause of the difference is increased use of context during a single fixation. The amount of context used during a fixation is embraced by the "effective visual field" which can be estimated by degrees of visual angle, typespaces, syllables, and words. Use of context is affected by nature of the task, subjective criteria, and factors within print materials. A group of 30 fast and slow adults (19-25) and a group of 48 fast and slow children (10.5-11.5) comprised the experimental groups. The method of investigation involved simulation of two successive fixations over a continuous sequence of words with sequential constraints manipulated. The amount read from the second fixation was the dependent variable. It was concluded that whatever other factors are operative, a better use of context enables a greater amount of graphical information to be used in individual fixations thus reducing the number of fixations needed to cover the same materials. Increased contextual information facilitates the use of visual information as well as effecting response bias in word recognition and also appears in so doing to produce spare capacity for visual processing for use at the more peripheral locations of the visual field.

Following somewhat the same line of investigation as Marcel, Jackson and McClelland (1975) tested fast and average college readers, scoring above 70% on a carefully constructed comprehension test, on four tasks (free report, forced choice, unrelated letters, peripheral span) to determine if reading speed is primarily sensory or cognitive in nature. Their results indicate that faster readers are capable of encoding more information from each fixation suggestion increased cognitive processing. The availability of more information facilitates the use of contextual constraints resulting in more efficient use of partial information. At the same time, the greater efficiency in encoding of visual information frees limited processing resources to deal with the conceptual content of what is read and guide the eye to a useful place to begin the process anew.

Rayner (1975) suggested that previous tasks used in determining the size of the perceptual span were either quite unlike normal reading or interferred with a reader's normal reading pattern and thus resulted in differing estimates of the size of the perceptual span. Recent on-line computer techniques involving eye-tracking systems also have limitations although these more nearly approximate normal reading. Ten undergraduates were exposed to CRT and on-line techniques which recorded the duration and the location of each fixation in reading and the movement time for each saccade. Results suggest the reader is able to make a semantic interpretation of a word that begins 1-6 character spaces from his fixation point; from 7-16 spaces the reader picks up gross visual characteristics such as word shape and initial and final letters. It seems clear that the capability of distinguishing between a word and a non-word, or making a semantic interpretation, is operative only within the area of high visual acuity–a rather small range for the effective stimulus. Other characteristics can be acquired further into the periphery to a maximum of five degrees of visual angle to the right of the fixation point. The overall results suggest slightly larger perceptual span and peripheral areas than reported in previous research.

Edfeldt (1975) examined the old question of whether or not it is possible to read during saccadic movements. Current theory is reviewed which suggests that reading (seeing) takes place during fixation and that the reader is functionally blind during saccadic movements.

Using newly developed apparatus to assess saccadic seeing, Edfeldt suggests that the good reader utilizes the ability to read and comprehend even whole words during saccadic movements thus obtaining more exposure time to materials than the poor reader. It is still to be determined if saccadic seeing is ammenable to training or if it is part of the individual's Central Nervous System blueprint.

Factors Affecting Rate and Flexibility

McConkie and Meyer (1974) suggest that task structure variables have a substantial effect on the behavior which subjects exhibit during a reading situation. Using 70 undergraduate subjects, and passages with MC questions built on higher level factors (Bloom), they investigated the influence of four variables (payoff structure, form of payoff structure, type of payoff, presence or absence of feedback) on reading rate and comprehension. The existence of a payoff structure, in and by itself, had little effect but the form in the structure produced significant changes in reading rate. Feedback appears to be a critical variable in producing reading strategy change.

Samuels and Dahl (1975) used 1,100 fourth graders and 84 college students in a modified Latin square design to determine the effects of establishing appropriate purposes on flexibility and reading rate. Subjects read a prose passage and answered detail and general overview questions. Repeated measures ANOVA revealed significant differences in reading speed when subjects knew beforehand the kinds and difficulty of the questions they would be asked; both groups scored significantly faster for overview than for detail questions.

Sullivan (1975) analyzed pre and post treatment scores on an adapted version of the Reading Adequacy READ Test for a one-week experimental flexibility module using university undergraduates and a control group. ANOVA and correlation revealed no significant relationship between rate and reading materials, factors of general reading ability and reading flexibility.

Brandt (1975) investigated the effect of two reading rate instructional approaches (SRA mechanical versus motivated), using Ohio State university students participating in a reading improvement course, in relation to internal/external (Rotter I-E) locus of control orientations. Single factor ANOVA revealed groups receiving motivated training increased rate significantly more than the non-motivated treatment. The correlational analysis between personality type and reading performance was significant only for the control group.

Coke (1974) evaluated the usefulness of reading rate as a measure of difficulty in two studies relating to text-derived measures of readability. Oral reading rate was affected by readability since subjects read at a standard syllable rate. When subjects rated each passage for comprehensibility, correlation of comprehensibility with readability indicated that subjects were sensitive to readability, suggesting there are limitations in using reading rate as a measure of reading difficulty.

Carver (1976) studied the functional relationship between word length, passage difficulty and reading rate using 300 passages sampled from cirriculum materials ranging from grade one through graduate level. Passage difficulty was measured by the Rauding scale. Reading rate of 83 college subjects decreased in relation to passage difficulty, but, when measured in standard words per minute, reading rate was constant. The results are discussed in relation to current theory relating redundancy to reading rate and eye movements.

Wood (1976) assessed the comprehension of general and study materials presented at various accelerated speech speed rates to 208 post-secondary technical students stratified by ability. ANOVA and post hoc comparisons revealed significant differences by reading ability in comprehending main ideas, details and conclusions, and by rate for details. Seventy-five percent of the subjects preferred a compressed speech rate of 200-250 words per minute.

In applying the Modified Packard Method of discourse analysis to recalls of readers trained in the Evelyn Wood method, Hansen (1976) found varied differences between slow and rapid readers. when time but not amount of material was held constant, slow readers showed significant gains in recall of operations; with time and amount of material held constant, rapid readers showed significantly greater gains on total operations.

Kamm, Otto and Dulin (1974) described development of a model for evaluating certain aspects of workbooks which should be considered when developing or selecting texts to be used in improving rate of comprehension. Factors include difficulty of material, difficulty of questions, background of the intended audience, and the reader's purpose.

Readability and Materials/Textbooks

Readability

Miller (1974) compared the predictive performance of the Flesch (multiple-choice-based) and the Bormuth (cloze-based) readability formulas using the McCall-Crabbs Standard Lessons in Reading and the 1967 Miller-Coleman Readability Scale as criterion measures. Findings confirm the contention of cloze proponents that cloze-based formulas generate higher validity coefficients on cloze materials than multiple-choice-based formulas generate on multiple-choice materials.

However, Miller suggests that results may be due to the inherent nature and mechanisms of the scores generated by cloze procedures rather than demonstrating the intrinsic superiority of the cloze. Liau et al. (1976) reported the development of four readability formulas based on cloze procedures which expand upon the earlier work of Coleman. The formulas, which account for an even higher percentage of the variance, were derived by choosing a more reliable criterion and a better selection of predictors selected according to ease of scoring as well as percentage of variance explained. A cross-validation study supported the development of the four formulas. Coleman and Liau (1975) maintain that there is no need to estimate syllables for computer readability analysis since word length in letters is a better predictor of readability than word length in syllables. Cloze was used to calibrate criterion materials. A new readability formula was generated with letters per 100 words and sentences per 100 words as the predictors. Both can be economically estimated with optical scanning equipment. Estimated grade level equivalents for cloze scores are also presented. McNinch et al. (1974) tested the effects of varying deletion schemes in cloze estimates of college texts in science, mathematics and social studies using 184 college freshmen from a psychology course. Deletion schemes should be adjusted according to discipline studied; science should have a low deletion ratio while social studies and English can tolerate a more moderate (every seventh word) ratio. A moderate deletion ration was also tentatively suggested for math materials.

Using six analysts, Thomas, Hartley and Kincaid (1975) estimated test-retest and interanalyst reliabilities for the Automated Readability Index, the Flesch Reading Ease Score and the Fog Count. All coefficients, with the exception of one Flesch measure, were above .94. ANOVA applied to measured working times indicated that the Flesch takes significantly longer to use than either the ARI or the Fog. Blumenfeld, Black and Morris (1975) investigated the interrater reliability of the Gunning readability index. The correlational analysis suggested that while the results of the two independent ratings were not identical, the interrater reliability was of sufficient magnitude to be considered adequate. The means exhibited a slight constant error suggesting rater training to decrease interrater error. Blumenfeld and Justin (1975) used fifty random samples of reading material and obtained highly comparable readability ratings by Flesch and Gunning indices.

Patterson (1974) analyzed the relationship of selected intersentence factors to cloze comprehension scores of passages from different subject-matter areas to determine if intersentence factors add predictive power to the Dale-Chall readability formula when cloze scores serve as the measure of comprehension difficulty. Cloze tests (seventh word deletion) were constructed and Carroll's and Yule's type-token formulas used to determine the diversity of lexical units and words in the passages. The verb/adjective ratio and the noun-verb/adjective-adverb ratio were used as measures of passage natural language content. Correlation, stepwise multiple regression and ANOVA revealed that diversity of lexical units as measured by Carroll's and Yule's formula, and diversity of nouns as measured by Carroll's formula, were significantly related to cloze comprehension scores. For these subjects and materials, new readability formulae emphasizing lexical density should be considered.

The cloze performance of 120 community college freshmen was analyzed by McWhorter (1974) to determine the influence of passage organizational structure on comprehension and two estimates of readability. Four passages representing time order, cause-effect, comparison-contrast and statement-support organization were used. The Bormuth and the Dale-Chall readability formulas were applied to the passages and a fifty item cloze test was constructed for each passage. Mixed results indicated the Bormuth formula provided a significantly higher estimate of combined passage difficulty than did cloze scores; observed cloze means were significantly higher than predicted cloze means for all but the time order passage; the time order passage was significantly more difficult than the other passages; and the Dale-Chall estimate of combined passage difficulty was significantly higher than the Bormuth estimate.

Morrow (1974) rewrote college level expository prose to determine if easier readability would aid comprehension of low achieving community college students. Material was adjusted by reducing difficult vocabulary, shortening sentences, and a combination of the two. Cloze tests were constructed for the original and rewritten passages and materials presented to three experimental and one control group of students from remedial reading classes. The findings support previous research in demonstrating that reduced readability aids comprehension. English ACT scores, I.Q., and Vocabulary and Comprehension scores had significant curvilinear relations with the cloze criterion scores.

Bankston (1975) applied Kane's readability formula for mathematics material and rewrote one mathematics lesson at three levels of difficulty. When presented to remedial mathematics students, the group receiving the lowest readability version had significantly better post test achievement than the other two groups.

Bassin and Martin (1976) and Bassin (1976) tested the effects of three redundancy reduction techniques on reading comprehension, reading time and rate. A newsarticle was reduced 10%, 30% and 50% by one of three reduction techniques: word frequency, grammatical, and subjective. ANOVA and post hoc comparisons indicated reduction method had no significant effect on comprehension except for the 50% level where the subjective method produced better reading performance. There was a decrease in speed in reading the reduced versions of text. Herndon (1976) examined the role of redundancy in textual processing and, using methodology from computer science and education, generated a system for computer control of redundancy.

Klare (1975) evaluated 36 experimental studies which examined the effect of readability of materials upon comprehension and/or retention. In each of the studies reviewed, materials were selected, the readability altered, the material presented to subjects, and the effect on comprehension and/or retention estimated. In 19 studies, altering the readability of the material had a significant effect on comprehension and/or retention; in 6 studies the effect was mixed; and in 11 studies the effects were nonsignificant. Comparisons among studies were made to elucidate factors which interact with readability including reading performance measures, the test situation, reader motivation, readability considerations in materials, content considerations, and reader competencies. Attention to such factors could have helped to reduce or increase the probability of significant results. Klare concluded that many writers succeeded in writing materials readable enough for readers to increase their comprehension and/or retention but the exact variables operating in the process are not clear. Tables are presented listing the major variables thought to have an influence on readability.

Bostian (1976) used three groups of 95 university students to test the effects of line width on reading speed and comprehension in brochures and booklets. ANOVA revealed one-column material to be significantly superior to two-column material for both speed and comprehension, contrary to previous legibility research findings.

Based on the assumption that existing readability indices lack validity in accurately judging the concept or ideational difficulty of passages, Carver (1975-76a) developed the Rauding Scale of Prose Difficulty (encompasses reading and understanding) as a tool for using subjective judgments to scale difficulty of passages. The RS consists of six anchor passages (from Bormuth, 1969) judged to be at the grade 2, 5, 8, 11, 14, and 17 levels of difficulty. New material is compared to the anchor passages and the opinions of three "experts" (achieved the criterion of judging five passages correctly) averaged to obtain a rating. Evidence from three studies is presented to support the reliability and validity of the method including: psychometric and interrater reliabilities; validity coefficients in predicting Bormuth grade levels; using RS to estimate grade level of reading materials commonly used at various levels; and, prediction of passage difficulty using the National Reading Standards Test developed by the author. Kibby (1975-76) subsequently pointed out that the idea of an anchored set of judgmental passages is a sound one but that the validity of the Rauding Scale is questionable on a number of points including: lack of connection with previous readability research; poor definition and lack of a theoretical rationale for idea or concept difficulty; grade level discrepancies in the anchor passages; problems with rater underestimation of passage difficulty; faulty study of sample reading materials; and, inadequate sampling, poor control, skimpy or inaccurate interpretations of data in the three studies. Kibby concludes that the scale adds nothing to our understanding of comprehension, it is not supported as a readability measure, and it provides no clear guidelines for rewriting materials. In a response to Kibby, Carver (1975-76b) presents detailed counterarguments on the points raised and contends that, despite the criticisms, the Rauding Scale is able to assign grade levels with substantial validity. The RS seems to offer the best "mousetrap" for estimating concept difficulty of passages at the present time.

Materials/Textbooks

Haskins (1975) described the development and assessed the concept validity of a technique for pretesting interest in printed messages. A number of studies are described which support the position that use of short descriptive verbal concepts is predictive of interest in the whole message. Haskins further discussed the technique as an efficient and inexpensive alternative to complete curriculum review.

Booher (1975) compared the relative comprehensibility of pictorial information and printed words in instructions. Six picture-word formats were examined (print only, pictorial only, pictorial related print, print related pictorial, pictorial redundant print, print redundant pictorial). Results, using 90 Navy enlisted personnel, showed pictorial information important for speed but print information necessary for accuracy with comprehension most efficient with pictorial related print and pictorial redundant print formats. Kammann (1975) tested the relative comprehensibility of flowcharts versus standard printed instructions using 90 scientists and engineers, 90 suburban housewives, and Bell telephone employees in a follow-up study. Two different flowchart formats were compared with standard paragraph instructions and were found to be superior in comprehension, accuracy and speed. Comprehension for printed directions appear to follow a general rule of thumb: directions are understood correctly about two-thirds of the time.

A number of studies analyzed the difficulty level of college textbooks in relation to the reading ability of students using them. Smith (1975) applied the Flesch formula to 162 one hundred word samples from nine accounting tests (3 high school, 3 vocational and 3 college level). ANOVA and post hoc tests revealed differences by level and for level by publisher but not for publisher alone. Mean readability scores increased in difficulty by level and were considered below the grade level for students for whom the texts were written. Nilagupta (1975), using a sample of 213 university students (Bangkok) enrolled in English as a second language reading classes, found a relationship between reading ability (Nelson and an informal test) and text readability (Flesch). The four textbooks had readability levels above the mean grade equivalent comprehension scores of students. By contrast, Slovak (1976) found no relationship between the reading ability (NDRT) of 760 community college students' comprehensive of texts scaled by readability level (Dale-Chall, Flesch). Kurzman (1974) found that the reading ability of open-admissions college freshmen averaged 10.5 while the readability of their textbooks averaged 14.0.

Working with adult materials Rakes (1973) assessed the readability of 29 units of material selected from books frequently used in ABE instruction using Dale-Chall, Fry and SMOG indices. The resulting readabilities were compared with difficulty judgments made by 43 ABE teachers with discrepancies noted between the formula levels and the teacher estimates.

Pyrczak (1976) found that 100 word passages of the instructions accompanying income tax returns yielded a readability of 9-10 using the Dale-Chall formula. He composed a set of comprehension questions on the standard instructions and a set for a revised simplified version. On the first test graduate students in education produced a mean score of 5.65–functionally illiterate. The simplified set of instructions produced a mean of 7.65.

Test Development, Evaluation and Use

Various techniques continue to be proposed for the measurement of reading behaviors; particularly in the area of comprehension of prose materials. The fundamental problem in such measurement has always been to develop techniques which enable one to estimate how much has been comprehended in reading and how accurately it has been comprehended. The inherent difficulty arises from the fact that the processes one is attempting to measure are internal and cannot be directly observed; they can only be inferred from observations of sampled overt behaviors that accompany the internal processes.

The research reviewed in this section is organized in two parts. First, studies are reviewed which report development of tests of various types. Then studies which assess characteristics and uses of existing tests are reviewed along with studies which discuss various desiderata relevant to psychometric properties of tests and their use.

Test Development

Anderson (1974) examined the possibility that cloze measures are a function of content achievement among adult learners. College juniors and seniors took cloze tests before and after studying a prose passage under five conditions: mc, reproduction passage cloze test, recognition passage cloze test, reproduction summary cloze test, and recognition summary cloze test. Post instruction mean scores were reliably higher than preinstruction scores in virtually all testing conditions lending support to the proposition that cloze measures can serve as indices of achievement.

Ohnmacht and Fleming (1974) replicated and expanded work on analyzing content effects and complexity of structure effects with three types of verb deletions and two types of noun deletions in cloze test development. Results with undergraduates generally confirmed earlier studies that suggest verb types do not lend themselves to categoric rank ordering from lowest to highest difficulty of structure. An additional finding concerned the "devastating effect" of the absence of post-deletion context on subjects' performance with all three verb classes and, to lesser degree, with noun classes.

Edwards (1974) examined the reliability and difficulty level of cloze tests of poetry using 600 university freshmen composition students. Results suggest: deletion pattern did not significantly affect reliability; no significant difference between reliability for cloze and for a Poetry Reading Test (STEP Reading Series II) on the same passage; choice of deletion pattern affected difficulty level, number of responses required did not affect reliability; and printed format had no effect on reliability.

Salup (1975) investigated cloze test performance and relationships between the Coop English Test–Reading and cloze scores using 776 university freshmen and reported correlations of cloze with reading comprehension of .36 and .63; KR20 reliability of .66 and KR 21 reliability of .41. Good readers had higher scores than on use of correct grammar and total cloze scores.

Lowry and Marr (1975) tested and supported the construct validity of Darnell's cloze entropy procedure.

Coy (1975) analyzed data for 38 functionally illiterate adults in assessing criterion and construct validity and reliability for an Adult Reading Rate Test. Correlations between the ARRT and reading gain scores were deemed insufficient (.60) for use of the test as a predictor of an individual's specific performance score; teacher judgment was inadequate as a measure of the rate of mastery of basic skills; factor analysis yielded only one fo the five ARRT traits hypothesized in constructing the test; and the reported reliability for the test was .81.

The development and field testing of an instrument deemed appropriate for the domain of on-the-job reading found to be essential in entry-level positions of selected plants of a large company was reported by Schoenfeldt et al. (1976). Reading skills found to be important in reading industrial materials were identified and related to the types of materials read on the job. A criterion referenced instrument was developed to sample the content domain with cloze procedure used to establish the criterion pass-fail cutoff in addition to the usual 75% (instructional) passing level. Dale-Chall readabilities were also computed for passages. Content validity procedures are evaluated.

A 50 sentence test of syntactic complexity was developed by Cox (1976) and validated using 21 ABE subjects and 26 college literates. Findings indicated that ABE students had more trouble than college subjects with the test.

Carver (1974) extends previously reported work and describes evaluation of the reading storage test by comparing it to a modified version of the cloze test in three experiments involving 48 college students reading passages at four levels of difficulty. In the RS testing rationale, the task is to determine whether a chunk of information contained in a passage was stored by providing part of the original material to see if this will trigger the essence of the passage. The RS test seemed to be more reliable and equally valid as cloze in measuring gain in understanding during reading. In a further study Carver (1974) compared results obtained with the RS test to those obtained when paraphrase tests were used in precisely the same circumstances. The experimental conditions replicated those used in the 1974 study. Based on the results, the author suggests the subjective paraphrase test could be an extremely sensitive test for measuring the primary effect of reading but it may not be reliable. On the other hand, the objective RS test appears to be less sensitive to the primary effects of reading but much more reliable.

One of the persisting problems in testing in reading is the establishment of an isomorphic relationship between the concepts and information contained in prose and the questions generated to measure comprehension of those concepts. The major problems in question generation are well known in college-adult reading. Earlier work by Bormuth and others described a technique for generating passage dependent questions, using a linguistic algorithm, that were judged by native speakers of English to be legitimate and nonanomalous. Finn (1974) explores this work further and argues that the use of the question generating algorithm can produce all possible questions in reading passages and also provides a measure of the amount of information contained therein. He concludes that the argument is substantiated and, in addition, it is not only possible to produce all possible questions but that it is also possible to specify the complexity of each unit of information in terms of the underlying sentences involved in the questions generated.

Evaluation and Use of Tests

Rozenkranz (1976) evaluated the efficacy of using cloze procedures in estimating the reading ability of 35 students in advanced ABE classes and 50 in high school equivalency classes. Using a 1/5 deletion ration on commercial ABE materials, high school equivalency materials, and passages previously tested on elementary and secondary students, cloze was found acceptable as a measuring technique for matching materials with students. Fewer than 25% of students found it uninteresting; only 7% of items were not attempted by at least 75% of the sample.

One problem in using cloze techniques, claim Carstens and McKeag (1975), is the fact that students may become discouraged when test results seem low in comparison with data feedback from more conventional assessment procedures. They analyzed pre and post cloze results for university juniors and found that different methods of handling scoring resulted in different rankings, an important consideration if such procedures are to be used in evaluating students.

Cochran (1974) found a correlation of .61 between scores on the Revised Ohio Literacy Test (a five minute test of sentence meaning) and the Wunderlic Personnel Test, using a sample of 631 normal adults. Both tests related significantly to school grades. Standard score norms are presented for the tests.

The validity of a priori scales developed to divide the Brown-Holtzman SSHA into two groupings–Study Habits, Study Attitudes–and a combined score of study orientation, were evaluated by Khan and Roberts (1975). Factor analysis results suggest that the a priori division into Delay Avoidance, Work Method, Teacher Approval, and Education Acceptance holds for the first three but not for the EA factor. The results are related to use of the test and the rationale for development of the original SSHA scales.

Ash and Stevens (1975) applied four readability formulas (Dale-Chall, Flesch, SMOG, FOG) to a structured position analysis questionaire (PAQ) consisting of 194 job elements of a worker oriented nature and found that the PAQ had a higher reading level than was appropriate based on the minimum education level of grades 10-12 recommended for respondents.

Hennessy and Loveless (1976) examined the psychometric properties of the Cooperative English Test, the California Reading Test and the Cooperative Guidance Placement Program battery with respect to their specific value as placement tests for students entering an open admissions community college. Factor analysis results for the 1,534 students revealed all reading tests loaded highly on the same factor, as did the grammar tests of the CET and the CGPP batteries.

Williams (1976) used five methods of measuring frustration to compare college student's degree of frustration on informal reading inventories. Use of a multitrait multimethod matrix indicated: no physiological variables could definitely discriminate frustrating from non frustrating reading material; difficulty level of materials appeared to affect comprehension more than pronunciation; Bett's criteria and self-rating scales tended to measure the same aspects of reading frustration and were moderately related; none of the methods appeared to definitely discriminate among frustrating and nonfrustrating reading passages or measure degree of frustration with a high degree of validity; and, Rose transformed GSR and heart rate were the most reliable of the five measures.

Hakstian and Cattell (1974) discuss results of their factor analysis of 57 variables from ability tests administered to 343 subjects ages 15-55. The report includes discussion of the first-order factoring only. The common-factor analysis revealed 19 intercorrelated primary abilities ranging from such previously well established constructs as Verbal and Numerical ability to less well-understood factors in the memory and fluency-originality areas.

Ohnmacht and Weiss (1976), using a computer simulated exemplar, considered various models of assessing construct validity and suggested that a third dimension, that of time of assessment, be added to the two dimensions currently being assessed in construct validity studies based on the supposition that learning to read is not a static proposition but a dynamic developmental process.

Pyrczak (1975) reported that results of a number of studies indicate a substantial number of items in standardized reading tests lack passage dependency and presents results from a brief study to confirm the observation. Preston's 1964 study was replicated by Keetz (1975) by having 98 freshmen answer reading test questions without having first read the text. The findings confirm the earlier work showing better than chance performance on questions and a positive correlation between test scores and general reading ability.

Word Recognition

Studies from the word recognition literature are reviewed which touch upon various topics related to older readers including: the role of phonemic, graphemic and semantic mediation in lexical access and word recognition; developmental changes in word recognition strategy; processing operations in word recognition; the role of units smaller than single words; and, the effect of various linguistic and affective attributes of words in lexical access.

A number of recent studies have examined the role that phonemic and graphemic codes play in accessing lexical meaning of words. Several studies suggest less skilled readers (poor decoders) may also be poorer comprehenders (Golinkoff and Rosinski, 1976; Perfetti and hogaboam, 1975; Golinkoff, 1975-76; Pace and Golinkoff, 1976; Levy, 1975) and that phonemic recoding is an extremely important skill for both younger and older readers. Studies also suggest a tolerance for weaker versions of phonological recoding and entertaining a dual access hypothesis with both graphemic and phonemic information playing a role in accessing lexical meaning (Baron and McKillop, 1975; Kleiman, 1975; Meyer, Schvaneveldt and Ruddy, 1974; Bradshaw, 1975). Skilled readers do tend to rely less on phonemic recoding.

Samuels, Begy and Chen (1975-76) compared the word recognition, speed and processing strategies of less able and more highly skilled children and adult subjects. Results indicated that more fluent readers were faster in word recognition, superior in ability to generate a target word given context and minimal cues from the target, and superior in awareness when a false recognition was made. The results are interpreted as support for the efficacy of training in word recognition strategies. The research of Eriksen and Eriksen (1974) confirmed the importance of first letter position and left-to-right processing order in word recognition and supported the fact that instantaneous decoding of the sound-spelling pattern of the language typifies proficient readers.

Green and Shallice (1976) asked whether phonological encoding is an obligatory stage in reading or whether direct access to a word's semantic representation is the general rule. With results from adult skilled readers, they argue that meaning may be accessed directly from visual form and does not normally involve prelexical phonological recoding; phonological effects occur after lexical access rather than before. They point to problems in the use of the lexical decision task in such studies. Kleiman (1975) suggests that phonological recoding may play an important role beyond the individual word level in printed discourse by providing the reader with a strategy for maintaining in memory the words of a sentence long enough for the sentence to be comprehended.

Fleming (1976) presented skilled adults with nonsense words in three groups in two studies to test the phonemic recoding versus the graphic encoding alternatives to lexical access.

Results suggest there are alternatives to phonemic recoding and that the graphic encoding hypothesis may be more appropriate. The important point is made that the coding process employed may be contingent upon the task required in various situations. Bradshaw (1975) suggests that the skilled reader of relatively simple material seems capable of extracting much of the semantic content of a passage without going through a necessary prior stage of phonological recoding and concludes that an interactive model of word recognition may best explain the process with perceptual units larger than the individual letter.

Mason (1976) compared the results of word-nonword recognition tasks and found that children and adults were alike in being affected by word familiarity and interaction between initial consonant frequency and word familiarity. Children were also affected by final consonant frequency and interaction between vowel regularity and word familiarity. Skilled readers can recognize four-letter words without extensive orthographical analysis. Unskilled readers may use the same processes but revert to phonological encoding when a word is unfamiliar and contains an irregular vowel.

Schwartz (1975) examined whether performance in recognition memory tasks depends on the recognition of a specific interpretation of a word (as originally encoded) or upon its orthographic-phonemic representation. A list of common words paired with weakly cued rare words was used. The data suggest the recognition of a word depends on its physical representation as well as its meaning. Under certain circumstances, phonemic and orthographic information may be as useful in a recognition task as the particular meaning of a word.

Rayner (1976) used a population of children and adults and demonstrated a clear developmental effect such that with increasing age subjects were more likely to choose alternative words with the same shape as the stimulus words in artificially constructed word trigrams, quadrigrams and quingrams. Heavy reliance was also made on the first letter in alternative choices through the fourth grade. The results support Gibson's hypothesis that there is a developmental change with increasing age in feature analysis and extraction. Kindergarten children do not have an efficient strategy in dealing with the delayed recognition task; second through fourth grade children look for alternatives with the same first letter; fifth and sixth graders continue this but also look for alternatives with the same shape; adults rely much more heavily on shape than first letter. Word shape, when it is conceptualized as being made up of distinctive features which make up the letters of the word, is an important variable in word recognition.

Lazerson (1974-75 found that the skilled reader of Modern English can tolerate noise in the featural display (line drawn through print). However, highly variable spelling has a more detrimental effect on measures of reading performance. Featural information serves as the catalyst which enables reading to occur but the skilled reader relies upon orthographic and syntactic-sementic information stored in the brain demonstrating that skilled reading performance is a highly cognitive process rather than being simply a phonemic or visual activity.

Santa (1976) attempted to determine whether whole words, single letters or spelling patterns operate in the recognition of isolated words. Using younger readers and adults, it was concluded that the development of reading maturity appears to correspond to more flexible word recognition strategies and an increased dependence on larger units of analysis.

Studies of the reading process have shown that skilled readers are able to extract the meaning of text without directed reference to the appearances or sounds of individual words; words can activate their logogens without phonological mediation and attention to words can also be minimal prior to their recognition. Underwood (1976) tested the effect of unattended printed words upon a picture-naming task to determine if print can be read even when it is not attended to. Findings suggest attention may not be necessary for word recognition provided that the thresholds of words are lowered by current contextual information. If attention is highly selective then only semantically related items will affect performance.

Forster and Bednall (1976) tested the assumption that lexical access involves a search process. In one experiment test items were classed as ambiguous or unambiguous and in another they were classed according to their syntactic properties. In both experiments, when the target of the search was a nonexistent entry, an exhaustive search was involved. Frequency of occurrence was no longer related to decision time as in lexical decision experiments. The search model adequately explains the procedure whereby the most common meaning of a homograph is accessed but the less common meaning is accessed in some completely different manner.

Ball, Wood, and Smith (1975) asked whether there are differences in the order in which the visual, acoustic-articulatory and semantic codes related to words are accessed when one reads a word. Previous work suggested that a semantic superiority effect exists in contrast to the notion that some visual processing must precede semantic coding. They concluded that the size of the identification units determine the order of processing with larger units being detected first. In perceptual search tasks, visual processing would occur first while in semantic search tasks the larger semantic targets are accessed first accounting for the semantic superiority effect.

Dallas and Merikle (1976) investigated the facilitating effect of semantic context on word recognition and concluded that when the word recognition process approximates the stages of stimulus recoding, lexical memory retrieval and response execution it does not seem reasonable to attribute the semantic context effect to the stimulus encoding or lexical memory retrieval states of the recognition process; semantic context appears to affect the late response execution stage of the word recognition process.

Karlin and Bower (1976) asked whether subjects could categorize a word semantically before they precisely identified the word itself and failed to support the notion. The problem of whether identification precedes classification is poorly formulated as of yet; the presence or absence of the category effect results from the interaction between the type of stimulus and the definition of the category decision for a particular task.

Rubin (1976) measured the probability of correctly guessing a missing word (every tenth word deletion) using four context conditions: all words before the missing word; all words after the missing word; all words before and one word after the missing word; and just one word after the missing word. The probability of correctly guessing a missing word was greater given the words before it than give the words after it. On a word-by-word level the advantage of forward context did not hold; providing one additional word after a missing word provides much greater constraint than providing ten additional words before it. Although the left-to-right advantage holds over the long range, in the short range stimuli need not be processed in their temporal order. The results are contrasted with previous views of positional redundancy drawn from communication theory.

Cosky (1976) addressed the question of whether we read letter by letter. If word recognition includes letter recognition, then the difficulty of a word corresponds to the difficulty of its letters. Letter difficulty was assessed in two discrimination tasks and a letter naming task and easy and difficult letter words exposed to adult subjects. Word frequency and length were also manipulated. There was no effect for letter difficulty suggesting that letter recognition does not play a role in word recognition and we do not read letter by letter. Whatever plays a role in word recognition is smaller than the word and correlated with word length in letters.

Travers (1975) reexamined previous data on the processing of words and letter strings and concluded the data added support to the hypothesis that visual features information from multiple letter positions within a word is normally encoded in parallel rather than undergoing a process of serial letter-by-letter coding. However, the data also suggested previous conclusions have to be qualified.

Sloboda (1976) presented readers with verbal and musical text which contained spelling and notational errors and found errors were least likely to be detected when they were in the middle of words or musical phrases. Inference of interior elements usually results from structural rather than visual functions.

Pearson and Kamil (1974) tested the time to process nouns and inflected and uninflected verbs and demonstrated word length to be an important factor in latency. However, the addition of inflections ranging from one letter to several added a constant to the response time. Further, the effect of both length and affixes vanished with practice. While the results suggest that the subject begins by using a strict letter-by-letter approach he begins to realize there are other strategies for recognizing words. Various hypotheses are tested leading to progressively better performance.

Farley and Yen (1976) analyzed the influence of affective properties of words on information processing time in a high-speed visual search task analogous to rapid skilled scanning in reading. Words extreme on emotionality yielded significantly longer reaction times than neutral words supporting previous research suggesting rated unpleasantness depresses learning.

Somekh (1976) had subjects write brief stories after exposure to a matrix in which neutral and emotive words were embedded. Independent judges were able to distinguish to a significant degree between the stories of subjects exposed to the emotive and neutral words under a brief exposure subliminal condition but not to a longer exposure supraliminal condition. The findings lend support to the notion that subliminal stimuli can be recovered in a variety of modes of experience—in this case symbolically transferred as imagery in a story.

The recall of adjectives and nouns from two prose passages differing only in imagery value of their adjectives was investigated by Hiscock (1976) to evaluate the effects of adjectival imagery on prose recall. Adjectives from the high-imagery version of the narrative were recalled more frequently than the low-imagery version. Adjective imagery had no effect on noun recall. An individual difference analysis revealed significant group by narrative version interaction visualizers outperforming verbalizers with high-imagery adjectives and verbalizers tending to out-perform visualizers when adjectives were low-imagery. The author suggests further research to define the extent to which adjectival imagery value affects retention of different components of prose, the mediational processes involved, and possible application to education.

Richardson (1976) investigated the effects of stimulus attributes upon word-recognition and found that imageability, concreteness and number of syllables did not affect performance; words with suffixes were no more difficult to pronounce than simple nouns; word frequency affected time for discrimination but not that for reading; number of letters in a word affected time taken to read it aloud but not the time taken to discriminate it from non-words; and, non-words which employed rules of English orthography and phonology were more difficult to discriminate from words than those which violated such rules. It was concluded reading aloud employs a grapheme-phoneme translation based upon a letter-by-letter translation of the stimulus; discriminating words from non-words obeying accepted rules is based on a wholistic analysis of the stimulus; and, discriminating words from non-words violating those rules employs a direct test of the regularity of the stimulus based upon the combinatory rules of English orthography.

Botwinick and Storandt (1974) compared a group of older and younger adults with respect to their qualitatively scored vocabulary responses after having been matched on the basis of their vocabulary responses scored in a traditional way. The results cast doubt on the generalization that vocabulary function is resistant to the ravages of time but psycomotor skills are not. Quantitative scores were correlated with qualitative scores and younger adults gave more frequent superior synonym responses suggesting a decline in qualitative aspects of vocabulary.

Wickelgren (1975) compared retention functions for children, young adults and elderly subjects in continuous recognition memory for words over retention intervals ranging from two minutes to two hours and found no significant differences in decay rate across the three groups supporting the hypothesis that memory storage dynamics are equivalent over different ages.

Hultsch and Craig (1976) investigated the hypothesis that older adults are more susceptible to recall inhibition than younger adults and also manipulated the degree of organization of the experimental materials. Organizational manipulation had no effect for either group and older adults were not as susceptible to response as younger adults.

Sentence Comprehension

Sentences form basic units of linguistic performance, and although a good deal of research has focused on detailing the techniques, mnemonics and cues that the language user employs in comprehending sentences in spoken and written form, the present state-of-the-art suggests that the development of a fully adequate model of sentence comprehension awaits further research and analysis.

Many questions arise relative to sentence comprehension. What are the contexts within which sentences are best understood? Does the reader have to apprehend the full structural description of a sentence for comprehension? What are the relative roles of syntax and semantics in sentence comprehension? What is the difference between recall of a sentence and understanding it? What aspects of deep structure influence sentence comprehension? Do sentences need to be fully decomposed into deep structure before they can be fully understood? What grammatical variables influence sentence processing? What order is followed in sentence processing? In addition to these questions, much research relates to methodological issues in measuring and analyzing sentence comprehension. The sections which follow present recent samples from this burgeoning literature to illustrate the theoretical, empirical and methodological activity taking place in sentence comprehension research which uses predominantly adult normal populations as subjects.

Many schemes have been suggested for the order of processing of linguistic information in sentence understanding based on the surface-deep structure descriptive paradigm in transformational generative theories of grammar which suggests the order to be surface structure to deep structure to semantic interpretation; the first step "breaking the code" of organization and structure and the subsequent steps breaking the code of meaning. Muma, Perry and Gallagher (1974) suggest that there may be some construct validity for the alternative notion that a semantic strategy has priority over a linguistic strategy in sentence perception and comprehension by adults. Differential changes in semantic-syntactic strategy may be developmental in nature with the evidence revealing a curvilinear function in which earlier stages (below 18-24 months) represent rudimentary semantic strategy followed by a syntactic strategy (18-24 to 48 months) and the early stages of an adult semantic strategy (above 48 months). A shift also occurs from iconic to representational cognitive-linguistic modes in middle childhood. Results suggest a semantic over a syntactic priority in adult sentence comprehension and point to this as a feasible area needing further study.

In examining the multistore model of information processing (sensory processing, long-short term memory et al.) Mistler-Lachman (1974) examined a levels of processing model for sentence comprehension which focuses on type of encoding rather than storage. Meaning recall and exact word recall were used to assess sentence memory and deeper comprehension was found to lead to better recall. The author contends the outcome supports the use of memory measures as estimates of comprehension. Implications for the multistore and the levels of processing model are discussed.

Much research has been conducted on how sentential material is stored in long-term memory but such research has not fully separated recognition and recall from comprehension. Carpenter (1974) presented a methodology that permits separate examination of sentence comprehension processes and subsequent sentence memory processes. Subjects engaged in a verification process where they had to read a sentence and then verify it as being true or false. The verification methodology was evaluated using comparative sentences.

Using the picture verification process of sentence comprehension, Carpenter and Just (1975) abstracted five processes of sentence comprehension which appeared to remain invariant across content and task. They argue that their model integrates sentence comprehension with the more general area of cognitive psychology. Tanenhaus, Carroll and Bever (1976) critiqued the Carpenter and Just model suggesting it suffers from theoretical arbitrariness in the choice of representation form and hold that the common set of assumptions that emerge from the models examined is inadequate as a base for either a theory of sentence comprehension or verification. In a later paper (1976), Carpenter and Just examine sentence verification models in light of the criticisms of the criticisms of Tanenhaus et al. and contrast three areas of disagreement between the linguistic based approach and an information processing approach (nature of comprehension, nature of sentence representation, subject strategies) and list errors in the Tanenhaus et al. paper. Catlin and Jones (1976) reexamined the data presented by Carpenter and Just and point out problems in the goodness of fit between the data and their theory. They suggest the fundamental problem is the failure of the model to account for the difference in falsification time between sentence-first and picture-first conditions reflecting the model's simplistic view of the inner-proposition comparison process: particularly as an account of sentence-picture comparison data for affirmative versus negative sentences.

The semantic component in reading is important. The reader is thought to extract the semantic content and transform it into its logical deep structure; in essence, we abstract out the semantic core (to long term memory) and discard the surface representation (short term memory). Kolers (1975) explored an alternative view and suggests that, in some instances (rapid reading in particular), the surface representation is an integral part of the encoded sentence. Speed in reading depends on pattern analysis operations directed at surface lexical representation. He distinguishes between operational or procedural memory and substantive or semantic memory and suggests in many cases procedural memory may be sufficient to the information encoded in contrast to the normal assumptions related to concern for memory of context.

Kurlee and Ramsburg (1974) underscore the increased attention being given to the role of semantics in sentence processing and suggest this is due to modifications in the theory of transformational grammar (generative semantics) and a revival of interest in attempting to define the concept of meaning. Using anomalous sentences, the time required to process a sentence over three types of sentences (equational, transitive, and intransitive) into which of three types of anomalies had been introduced was the criterion. Assumptions underlying relative difficulty and the perceptual clustering of phrases and meaningful units were supported by the findings. The authors had to add cognitive assumptions in order to account for the findings.

Theories of sentence comprehension deal largely with a single literal level of meaning. Harris (1976) reports a study which attempted to examine how comprehension of metaphors takes place on the assumption that it operates in a two-stage process; first understanding the words in a literal way and then using the literal meaning to construct the intended metaphorical interpretation. Evidence for the two-stage process has been found for other types of non-literal language. Using metaphorical interpretations from Shakespeare and non-metaphorical equivalents, no differences between the two sentence types were found, using a paraphrasing task, in latency and adequacy of the paraphrase. Thus, no evidence for the two stage hypothesis was obtained and metaphors do not require processing similar to that of other nonliteral language; they introduce no inordinate difficulties in comprehension in comparison to equivalent non-metaphorical interpretations.

Just and Carpenter (1976) explored distinctions between the processes that take place in sentence memory and sentence comprehension. The experimental task required subjects to read a sentence and verify whether it was true or false at initial sentence presentation. Results indicated that complex sentences are often recoded into a simpler form during the act of comprehension. Later during recall the recoded representation may be retrieved. Thus, the "meaning for gist" effect may be due largely to a "comprehension of gist" process.

Winter (1974) conducted four experiments in which college students read sentences 9-13 words in length after hearing four context conditions: word actually in sentence, cue to a word in a sentence, 30 word two sentence context either general or specific, no context. Time to process sentences was the dependent variable. No significant differences attributable to the context conditions were found, providing context did not facilitate processing of sentences contrary to conventional wisdom assumptions about the efficacy of context in deriving meaning.

Much research has been done in assessing the role of overall grammatical complexity and different types of grammatical units in sentence recall and comprehension. Layton and Simpson (1975) reported an experiment which investigated whether the match mis-match between active and passive sentences and questions would be observed when the comprehension task was made more difficult by increasing the load on memory. When one sentence was presented less error occurred when sentences and voice were the same than when they differed. Both surface and deep structure are retained after two questions; only the deep structure of sets of four and eight sentences are retained. The view that surface information is retained in memory when there is an appreciable memory load has to be questioned.

Cook (1975) examined the variety of explanations advanced for the difficulty of plurally imbedded relative clauses in English and advanced the hypothesis that errors in comprehension of relative clauses are caused by perceptual strategies resorted to when the normal capacity of the processing channel is exceeded. Results using native children, foreign adults and native adults suggest this to be the case. An overload is reached with a single imbedding for children and foreign adults and with a double imbedding for adult native speakers.

In exploring the grammatical dimension of reading it has been found that reading performance at all levels indicates application of the reader's implicit knowledge of the grammatical structure of the language. Sawyer (1975-76) found the type of grammatical structure being processed affects reader dependency upon cues in the visual display; more predictable phrases are recognized with only minimal visual cues while less predictable phrases require a greater number of details. Grammatical constraints operating within passive right-imbedded sentences offer firmer syntactic cues than active-left imbedded sentences. Agentive phrases and locators share an equal likelihood for occurrence as skilled readers pose hypotheses about what may be expected to follow a passive verb.

Although several theoretical positions and a variety of empirical tasks suggest the primacy of verbs in carrying the semantic core of meaning in a sentence, verbs are strikingly inferior in memorial tasks and nouns are generally recognized and recalled better. The verb of the sentence defines the plot in a play; the subject merely indicates one of the actors. However, when subjects know they must remember something about the play (sentence), it is the actors (nouns) that are remembered. This has become known as Fillenbaum's paradox. Reynolds and Flagg (1976) attempted to resolve the paradox by several experiments which explored the efficiency of various sentence elements in recognition memory. The total results suggest that verbs are processed to a deeper semantic level while nouns have the form (orthographic, phonological) information recorded more strongly. The theater metaphor holds and is supported by a semantic encoding model.

Sentences with self-imbedded relative clauses have received considerable attention in the past decade and are regarded as being extremely difficult to comprehend. Much research has gone into understanding why. Hakes, Evans and Brannon (1976) suggest that experiments purporting to demonstrate this have confounded the self-imbedding or right branching location of the relative clause with their internal structure, comparing self-imbedded objective relatives with right-branching subject relatives. Results show that when internal grammatical structure of the relatives is carefully controlled, there is no evidence whatsoever that sentences containing self-imbedded relatives are more difficult to comprehend than ones containing right-branching relatives. They suggest that findings of past research can be accounted for by the fact that object relative clauses are more difficult to comprehend than subject relative clauses rather than self-imbedded relatives being more difficult than right-branching relatives.

The skilled reader interacts with graphemic, pattern semantic-collocation and syntactic considerations beyond word identification in reading. Wisher (1976) explored the idea that our expectation of what is to come provides a catalyst that guides the mapping from grapheme to meaning in developing a hypothesis which is tested against the incoming data. With this information the reader assimilates what is expected with what is first seen. Younger and older readers do anticipate syntax, almost on a word by word basis. Wisher examined the influence of knowing the syntactic structure of a sentence beforehand in reading. He confirmed that such knowledge allows a sentence to be read with less effort and faster by reducing the need to store in short term memory and engage in linguistic computation. Anticipation of structure is an important skill in reading.

Kleiman (1975) explored the role of speech recoding (silent speech, inner speech, subvocalization) in sentence comprehension in three experiments and concluded that the results provide support for a model of reading sentences in which speech reading occurs after lexical access thus facilitating the temporary storage of words necessary for sentence comprehension.

Newstead (1976) questions the primacy of syntax in sentence processing and suggested that semantic constraints are at least as important as syntactic factors in sentence processing. Using reversible sentences (interchangeable subject object) in three experiments it was shown that semantic constraints exist and relate to sentence perception; the power of the constraints is related to sentence complexity; and the constraints derive from the overall relationship described in the sentence. The results of previous work which failed to find semantic constraints could be explained in terms of the strategies subjects adopted in the various tasks.

Holyoak (1974) re-examined previous research on semantic memory and the effect of imaging sentences on verification tasks by having subjects rate the semantic relatedness of subject and predicate words in sentences. It was concluded that the effects of rated imaging could not be distinguished from the effects of the semantic variables. Greater comprehensibility was reported for high-imagery sentences. Many questions remain on the role of imagery in sentence verification and research in this area can advance only if variables like semantic relatedness can be controlled.

Pezdek and Royer (1974) explored the proposition that concrete sentences are stored as visual images whereas abstract sentences are stored as verbatim strings. Poor detection of many subjects may be due, however, to their comprehension of sentences rather than differences in storage mechanisms for abstract-concrete sentences. Recognition for meaning changes in abstract sentences was significantly higher for sentences embedded in a paragraph designed to increase comprehension than for the group presented sentences without paragraphs. There was no appreciable differences for groups in recognition of wording changes in abstract sentences or in recognition for both wording and meaning changes in concrete sentences. Results are discussed in terms of different storage mechanisms for abstract and concrete sentences.

Davis and Proctor (1976) had subjects observe both concrete and abstract sentences prior to carrying out a distractor task which involved either verbal or perceptual coding. In agreement with their hypothesis, sentence type interacted significantly with task. Abstract sentence recall was significantly more impaired following the verbal task than the perceptual, while this effect was reversed for concrete sentences. In another experiment, this difference also reached significance, and further, error analysis confirmed that perceptual task influenced the semantic context of the sentences recalled. The studies lend support to the contention that the mediating processes underlying the apprehension and retention of concrete and abstract sentences are qualitatively different. Mental imagery plays a role in internalization of concrete materials but abstract sentences are mediated by an essentially verbal process.

Learning From Prose

In recent years there has been a steady increase in the amount of research reported on learning from prose. As Frase (1972) points out, "The cumulative record of research on prose materials over the past ten years is a positively accelerating function" (p. 337). Much of this has been stimulated through the initial work of Ausubel on meaningful verbal learning and Rothkopf's notion of mathemagenic factors in text study.

With the focus of many college-adult reading programs being, in essence, learning to learn (Postman, 1969), studies which attempt to point out the importance of the manner in which prose materials are organized for presentation, and what the student brings to and does during the study of text, can provide fruitful insights and research directions in the examination of college-adult reading. Bower (1975), in analyzing story understanding and recall among college age populations concluded, "... I think the topic (processing and understanding text) is too central to cognitive psychology for us to continue postponing investigating it until we first figure out how lists of nonsense syllables are learned" (p. 533). The most recent NRC reviews (1974, 1975) included scattered reports on prose learning. Because of the increasing potential of this area of study, this review includes a good sampling of the more recent research produced as a result of the activities of educational psychologists examining basic notions in prose learning and mathemagenic activity.

A number of critiques and reviews related to prose learning have appeared: Frase (1968, 1970, 1972, 1973, 1975) and Carroll (1971) review earlier studies; statistical techniques employed in adjunct question studies were reacted to by Ladas (1973); Anderson (1970) and Anderson and Biddle (1975) analyzed the prodiguous literature on interspersed and inserted questions; and Rothkopf (1970, 1972) provides an overview of the mathemagenic concept. The recent review by Faw and Waller (1976) organizes the literature into useful categories (advance organizers, response modes, objectives, inserted questions) and provides an excellent overview and critique of prose learning research through early 1975. Although there is, inevitably, some overlap this review extends the Faw and Waller critique with subsequently published materials.

Notetaking

Assuming the hypothesis that the greater the degree of interaction one has with material and the greater the personal idiosyncratic organization one can place on materials to be learned the greater the recall, notetaking has always appeared to be a useful study skill. Howe (1974) reviewed research which examined learner strategies in producing and reproducing information made available in prose and the facilitative effects of notetaking. Research continues which attempts to provide support for the utility of notetaking and to define the optimal set of circumstances in which this skill is beneficial to learners.

Fisher and Harris (1973) compared several groups on retention of materials presented in lectures and noted no difference between groups on mc and free recall measures. In a later study (1974), they noted that subject's preference for form in taking notes may interact with the experimental condition; subjects appear to do better when they are allowed to encode in the way that they prefer. They also noted the lack of a significant review effect and superiority of females over males.

Peters (1972) varied the presentation rate (normal 146 wpm, fast 202 wpm) of lectures and found notetaking significantly reduced recall for both conditions on a mc test. Negative effects were most evident for low efficiency readers. Aiken, Thomas and Shennum (1975), however, found that students who took notes during a brief interval in which the lecture was stopped, recalled significantly more than students who took notes while the lecture was in progress or no notes at all.

Gilbert (1975) analyzed the effects of notetaking versus listening on connected and disconnected verbal material and concluded the physical act of notetaking need not interfere with the comprehension of spoken material; notetaking was not more effective in immediate recall of disconnected as versus connected discourse.

In a study assessing notetaking and review, Annis and Davis (1975) analyzed students assigned to seven treatments and a control group with notetaking conditions manipulated to test notetaking as an encoding process and notetaking as an external memory device for review. Results suggest that both encoding and external memory functions are important for successful recall but it is the act of personally encoding the lecture which provides the best retention. An external memory device is important for review, whether it is one's own notes or those provided by the lecturer. Mental review alone was singularly unsuccessful in retention. Fisher and Harris (1973) also examined the encoding and external memory aspect of notetaking. Analysis of three groups varying in amount of review revealed superior recall for those who reviewed their notes suggesting the external storage function of notetaking is a critical factor in review for recall. Results of a study by Carter and Van Matre (1975), analyzing the relationship of notetaking and review, supports the desirability of taking and reviewing notes for maximum retention and recall efficiency. They suggest the external memory function is of the greatest importance with encoding differences being minimal in importance.

Palmatier and Bennett (1974) examined the current perceptions of 223 university students on taking and reading notes in study and found students reported regular notetaking and view notetaking as important in study. No differences were found between successful and highly successful students (GPA); more successful students seemed to be better notetakers. No one appears to be teaching notetaking although such device advice consistently appears in how-to-study texts. As has been pointed out in studies of notetaking over the years, notetaking is a skill that must be learned to have maximum pay-off.

Dyer (1975) assessed the effects of notetaking on retention of printed materials. Both notetaking and rereading improved encoding and recall but the effects were not additive. Encoding efficacy was emphasized in notetaking from print. Also working with print materials, Rickards and August (1975) studied the effect of subject-generated and experimenter-provided underlining of sentences (one per paragraph) and found subject-generated groups recalled significantly more non-underlined material, supporting somewhat the encoding hypothesis advanced for notetaking. Idstein and Jenkins (1972) used differing passages of text and found that underlining was not significantly better than repetitive reading for either passage. Friedman and Wilson (1975) glued pages of course texts sold to students and examined them after the course to determine text usage and amount of underlining of major and minor points. Students made greater use of the earlier part of the text, the earlier parts of the chapters, and the chapter summaries; proportion of underlining followed a similar pattern with females underlining more, and no difference in text usage was found in use of major or minor points or between students taking essay and mc exams.

Faw and Waller (1976) observed that in notetaking studies time was usually controlled and control groups utilized, studies used aural rather than visual preserntations suggesting more work is needed with multiple presentation modes; and a need exists for studies assessing individual differences. Treatment by aptitude research could reveal more promising interactions and implications for practical application in this area.

Prefatory Objectives

Research on the effect of instructional objectives as orienting stimuli in prose learning continues to be reported.

Jenkins and Neisworth (1973) found that performance on test items was consistently superior when the corresponding objectives were provided than when they were absent. Duchastel and Brown (1974) reported that students provided with a list of objectives related to text learning did better than controls on objective relevant content and poorer than controls on objective irrelevant questions suggesting that objectives enhance learning and may depress incidental learning. Hesson (1975) generated significant correlations between use of performance objectives and basic reading skills, mastery of text material, and knowledge of objectives. The group receiving training in use of objectives scored higher than the group receiving objectives alone. LaPorte and Nath (1976) found that a hard specific goal prior to learning increased individual performance goals and correspondingly increased the information gained during reading.

Lawson (1974), Gagne and Rothkopf (1975), Rothkopf and Kaplan (1972), Kaplan (1974), Kaplan and Simmons (1974), Kaplan and Rothkopf (1974), Rothkopf and Billington (1975a) and Frase and Kreitzberg (1975) provide substantial evidence that learning is enhanced by objectives; intentional learning is facilitated at the expense of incidental learning, although the effect of incidental learning is less clear. Specific, distributed objectives enhance learning and may also serve a review function. Duchastel and Merrill (1973) caution that objectives may not always be effective; if they are too numerous they overwhelm and confuse the learner and in and by themselves rarely reduce study time. Duell (1975) noted no significant effects for specificity and knowledge of objectives on mc test performance and concluded that stating objectives may only enhance learning of materials that are considered unimportant by subjects. Stumpf (1974) found that prefatory objectives did not affect immediate achievement of community college students. When number of goals and similarity between the wording of goal descriptions and segments of text were systematically varied, Rothkopf and Billington (1975b) concluded that too many goals decreased learning and text specific objectives enhanced learning.

Faw and Waller's (1976) review supports the facilitating effect of objectives on intentional learning when objectives are more specific, less dense, and distributed with each section of the text. They point out that efficiency per unit of study time needs to be carefully considered in such research.

Advance Organizers, Orienting and Active Procedures

Advance organizers and orienting procedures in prose learning continue to interest researchers. Advance organizers consist of various kinds of introductory expositions presenting either new concepts under which detailed future learning can be subsumed or drawing distinctions that enable the learner to discriminate new concepts. Orienting procedures consist of learning instructions, activities, and information, beyond adjunct questions, which allow the individual to become actively involved with the learning materials. The activities which subjects carry out in memory processing are important.

Advance Organizers. Andrews (1975) compared posttest performance of subjects who heard either a cognitive organizer, a list of directed reading questions, or both prior to reading a brief passage. The cognitive organizer group demonstrated the best retention of the passage. Schnell (1973) studied groups of college students who were provided with preorganizers, postorganizers, both, and no organizers in learning from a passage on neural maturation. The three groups receiving organizers exceeded the control group with the postorganizer group performing best. Rickards (1975-76) organized five conditions: superordinate precontent (advance organizer), superordinate postcontent (post organizer), coordinate precontent, coordinate postcontent, and control to determine the processing effects of conceptually subordinate and coordinate organizers interspersed in text. Superordinate precontent statements lead to significantly more recall of organizers and subordinate facts than any other type of inserted statement. Advance organizers produced more contiguous recall of subordinate facts with organizers per organizer than did post organizers. Only advance organizers generated conceptual recall incidental to the organizers themselves.

Jenkins and Bausell (1976) pursued Ausubel's notion that the learner's present knowledge is important in determining subsequent learning. Four experiments were reported. In the first prose learning was facilitated when meaning of key terms was built into cognitive structure. The second experiment was run to determine if the results would generalize to school settings and produced discrepant results. Experiments three and four further analyzed the effect of number of terms and ratio of terms to prose material. The effect of number of terms was disconfirmed and the effect of the ratio of terms to prose material confirmed. Relevant ideas anchored in cognitive structure appear to facilitate learning only in materials with relatively few contextual anchors and establishing anchoring ideas in cognitive structure prior to learning does not necessarily transfer to fairly liberal school-like conditions.

Faw and Waller (1976) enumerate many of the problems in advance organizer research, related to lack of definition and the real effects of pre and post organizers, and conclude that when materials are difficult or novel, or when subjects' ability is limited, organizers that are general or abstract can facilitate retention of materials. Studies reported to date are limited methodologically and, although advance organizers may prove to be beneficial in learning from prose materials, the effect has not been conclusively established. For example, Schumacher, Liebert, and Fass (1975) explored the interaction of the amount of higher level structural information given to subjects (advance organizer) and passage organization. A passage typical of textbook conditions was presented to 144 college students in either of two formats: one long paragraph or six separate paragraphs. One group received an advance organizer in reading the passage. ANOVA revealed that subjects performed better under the whole rather than the part paragraph organization. Contrary to expectation, the interaction effect showed that the paragraph organization main effect was due primarily to differences between the two paragraphs when an advance organizer was not present. The passage with neither paragraph cues nor an advance organizer produced better recall than the passage with paragraph cues but no advance organizer. The advance organizer passage was intermediate regardless of paragraph structure. The authors hypothesized that subjects, in the absence of advance organizers, carried out activities aimed at generating their own structure for the passage. A subject-generated plan for a passage may be superior to an experimenter-generated plan.

Orienting and Active Procedures. Arkes, Schumacher and Gardner (1976) studied the effects of several orienting tasks on retention of prose materials. Three semantic tasks were employed:

copying, outlining, and sorting constituent sentences, and two nonsemantic tasks, circle e's and es's in a passages. Recall and recall per unit time (efficiency) were dependent variables. Outlining and copying tasks generated good recall by forcing interaction with the material. Non-semantic tasks scored lowest with the sorting task affecting the quality of the interaction with the material. Strong instructions induced more interaction for intentional conditions. Maximum recall and efficiency occurred with simple instructions to read the passage with longer time periods leading to better recall.

Frase and Kreitzberg (1975) analyzed the effect of telling readers what sentences to learn by referring to words in sentences or topics of sentences. Five types of learning directions were used: word-positive, word-negative, topical positive, topical negative, and control (learn the entire passage). Word references but not topical reference facilitated relevant recall and depressed incidental recall. The results support the importance of precise references in the use of learning directions. One cannot assume, however, that active response invariably leads to increased learning. Howe and Singer (1975) asked subjects to read a passage of prose under different conditions chosen to represent realistic study strategies. In analyzing three conditions (reading, copying, summarizing), contrary to expectations, reading alone was superior to both copying and summarizing with summarizers learning more than copiers. Results held for both long term and immediate retention tests. In a subsequent experiment, reading the passage was contrasted with listening to the passage involving six different combinations of presentation and subject activity. The findings of the first experiment were replicated but, with auditory presentation at dictation speed, there was no significant difference between subjects who read and subjects who recorded. The authors suggest that the interaction of presentation forms be considered in research on meaningful prose learning. Fronk (1974) reported that the listening mode was superior in recall for all students when using materials that required physical manipulation. In units requiring no physical manipulation, listening was more difficult for poor readers only.

Richards (1975) sought to determine if large numbers of college students could modify their behavior via a combination of study skills advice and behavioral self-control techniques. Treatments were delivered through typed handouts. Self-monitoring (observing and recording own behavior) was found to be an effective addition to study skills advice but stimulus control was not. Stimulus control may be too complicated and unwieldly to teach in large groups, may not be ammendable to typed handout treatment, and may require too radical a change in student behavior. The efficacy of the handout treatment for a large group of subjects is an interesting finding in this study and confirms the results of Beneke and Harris (1972) who found handouts worked as well as group therapy seminars in teaching study behaviors.

Ross and DiVesta (1976) examined the effects on retention of listening to versus actually presenting oral summaries of prose materials. a 25% advantage in recall of relevant concepts and facts was found when an oral review was required compared to no review, and a 13% advantage was found in recall when an oral report was made compared to listening to a review. Oral review is an effective study strategy for enhancing recall of meaningful text material. Oral activities produced positive significant low correlations with anxiety symptoms suggesting anxiety serves an arousal function relative to learning but it need not be debilitating in terms of performance.

Crouse (1974) investigated the acquisition of course material under repeated testing and found: with limited study intervals, acquisition increased over trials of study followed by testing; with student determined study intervals, no effect was apparent for repeated testing. However, repeated testing in the present study was important primarily as a way of increading study time. If study time can be increased by other means, repeated testing may have no advantage.

Questions

The effect of questions inserted in prose learning has been attributed to the control which they exert over the general classes of mathemagenic responses, taking place in inspection and processing, which the learner gives to prose learning. Research in this area continues examination of such variables as question position, type of question, passage characteristics, source of questions, immediate and delayed recall, incidental and intentional learning, reported study time, and learner characteristics related to prose learning.

Meader (1975) utilized multiple matrix sampling techniques to cover all questions subdomains (Bloom's Comprehension and Knowledge categories), and Bormuth's procedures to insure production of test questions most relevant to prose articles, in testing the effect of inserted questions and generated questions. The inserted post question group performed significantly better on knowledge and comprehension than either the generated questions or control group on immediate recall, but not on the delayed retention test. Packwood (1976) inserted questions before and after prose passages and found significant main effects for topicality but not for position or spatial contiguity. A significant position by topicality interaction was noted suggesting pre questions positively affect retention while post questions do not. Hiller (1974) reported that looking back does not necessarily facilitate learning but this may have been due to the control of study time. Boker (1974) demonstrated that inserted questions have the same effects when assessed by delayed and immediate tests of retention.

Snowman and Cunningham (1975) examined the effect of pictorial and written questions interspersed through text in four conditions. Reader generated pictures and experimentor provided questions were equally facilitative supporting the notion that pictures have an effect similar to that of questions and suggesting that the range of study of adjunct aids in prose learning be expanded to include visual and verbal prompts. Royer and Cable (1975) suggested that subjects reading an initial concrete passage followed by exposure to an abstract passage would have a new knowledge structure in contrast to subjects who first read an abstract then a concrete passage. The concrete first subjects recalled a minimum of 40% more material. In a subsequent study (1976), they compared the effects of passages labelled abstract with illustration, abstract with analogies, concrete, unembellished abstract, and control. Results provided substantial support for the contention that the first three passages would have a greater facilitating effect on learning.

Rickards (1976a) found that the generally noted postquestion superiority effect did not hold for all type of verbatim adjunct questions. Verbatim postquestions did not generate incidental learning. High structural importance postquestions did not generate subsumptive activity but they only required literal recall of the high level information they solicited. When deeper text processing is required, significant incidental learning has been found to be related to question type. Rickards (1976b, c) subsequently examined the effect of conceptual and verbatim questions placed prior to and after text segments. Conceptual prequestions were superior to conceptual postquestions yielded less recall than verbatim postquestions. Only conceptual questions were superior on delayed recall, they produce more highly structured and organized memories than verbatim postquestions.

Facilitative effects for higher order conceptual questions have been found in a number of studies (Frase, 1973; Shavelson, et al., 1974; Rickards and DiVesta, 1974; Felkar and Dapra, 1975; Mayer, 1975), although the effect can vary according to learner characteristics, question position, type of learning, nature of the prose material, and frequency of questions.

Adejumo (1975) compared the results of self generated questions, questions generated by other students, experimenter generated questions, and no questions finding a significant omnibus F. Post hoc comparisons revealed that the group using questions generated by other students performed significantly better than the self generating question group for total and incidental learning and for intentional items. The experimenter question group performed significantly better than all others on intentional learning. GPA also differentiated subject performance. Morse (1976) reported groups who used questions, either pre or post or experimenter or reader generated, did not differ significantly on the posttest from the control group but experimenter generated questions tended to be more facilitative than reader generated questions.

Voss (1974) analyzed effects of two types of post questions (mc and completion) and found that performance failed to improve over testing trials. Multiple choice questions, with repeated use, actually produced interferences with higher level learning. Rickards et al. (1975-76) reported that, with respect to intentional learning, inserted number and common word questions yielded equivalent results regardless of text position. Only common word postquestions, however, produced significant recall results. Neither common word nor number postquestions, however, produced significant incidental learning results with number questions producing a slightly negative mathemagenic effect. Hiller (1974) examined the effect of easy and difficult distractors finding that hard questions require more processing.

Studies by Howe and Colley (1976a, b) suggest that prose learning will be influenced by the set established by passage variables and the nature of inserted questions. Appreciably better performance was reported when questions were similar to those asked in earlier passages but not when earlier and later passages were similar.

Rothkopf and Billington (1974) found that performance on a test item is facilitated when there is good veridicality between the test item and the text content; both when reinspection is allowed and forbidden. In addition, attempts to answer a question tends to improve subsequent ability to answer other questions from the same topicospatial neighborhood of the text, but this indirect priming effect was small.

King (1974) suggested that total learning time may be the most important variable in connected discourse learning and Faw and Waller (1976) found lack of control of the study time variable to characterize much earlier research in this area. More recent research is beginning to examine this factor. Bull (1973) and Bull and Disney (1973) point out that the primary value of questions may be their arousal effect. However, Zilman and Cantor (1973) reported that questions did not arouse curiosity but suggested the arousal function may be most important when attention is failing rather than in first learning.

As pointed out by Faw and Waller (1976), the major issues in the mathemagenic effect of questions relate to verbatim recall versus comprehension (Anderson, 1972), applicability to normal classroom learning situations with extended textbook materials, the extent to which different ability levels respond to inserted questions, and incidental as compared to intentional learning effects. Only recently have individual difference variables been linked to question use (Shavelson et al., 1974; Adejumo, 1975; Sanders, 1973). Hiller (1974) found learning from prose text to be correlated with anxiety and self-confidence of subjects. Koran and Koran (1975) reported that frequent questions may be best for low ability students through reducing the burden of semantic processing and strengthening associative learning. Their analysis of instructional treatment and learner characteristic interactions revealed highly significant treatment effects on the ANOVA omnibus F. Optimal question pacing varies according to learner characteristics with differences found on incidental but not intentional learning. ATI studies (Salomon, 1972) could provide explanatory principles but we are still some distance from developing decision rules which will allow for the assignment of students to alternative treatments designed to reach the same terminal objectives (see Cronbach and Snow, 1977).

One of the broadest applications of mathemagenic questions in learning from text was reported by Larkin and Reif (1976). Introductory college physics students were taught a skill for acquiring from a text description of a quantitative relation a generally specified set of abilities which are prerequisite to applying relations in problem solving. To develop the skill, students studied text with the interspersed questions requiring demonstration of the specified abilities. Such training significantly enhanced acquisition of these abilities from a subsequently presented text description of new relations, demonstrating transfer effect. A person's strategies for learning from text can be modified through use of systematically designed study materials.

Text Structure

The organization of prose can have substantial effects on what the student remembers from reading (Rothkopf, 1972). Our knowledge about the processes involved in remembering from prose learning lags far behind that about memory for word lists. Several recent representative studies which examined variables in text structure affecting the complex process of prose memory in college-adult populations are presented here.

Steingart (1975) tested the effects of organization by referrent, attribute and random organization using mc questions and written recall. Half the college subjects also visualized the pages. Referrent, attribute and random organization of text produced the best comprehension in that order. Regardless of text organization, imagery instructed subjects stated significantly more text relations; imagery also resluted in a significantly higher level of inference making and recall protocol organization which was consistent with figural representations of text information. Imagery functions to form and preserve unitized representations of text information.

Siegal, Lautman and Burkett (1974) conducted three experiments in which text readability was lowered and tape recorded supplements were made to text to assess the effect of text modification and audio supplementation on text comprehensibility. No gains were evidenced through lowered readability (Air Force technical training materials and look-up manuals) or auditory supplementation. However, for home study, as compared to on-the-job instruction, lowered readability resulted in increased criterion scores.

Using Cromer's difference model, Johnson (1975) classified 200 college freshmen as Good (average or above average comprehension), Difference Poor (average or above average vocabulary but below average comprehension), and Deficit Poor (below average vocabulary and comprehension) readers to examine the interaction between readability, subjective organization ability and passage organization. ANOVA on literal and inferential comprehension post scores revealed differential effects supporting the hypothesis that readers use text organization processes in learning as well as reading for comprehension. The activity approximates an information processing paradigm.

Comprehending prose is not based on literal recall of exact words and research continues on the subunits which are most likely to be remembered in prose learning. Johnson (1974) found the patterning of recall of linguistic subunits of intact textbook prose was strongly related to semantic dimensions of abstract/concreteness, specificity of denotation, comprehensibility, and interest on tests of immediate and delayed recall. Presumably, modification of texts on the four factors would result in increased recall.

Brown (1975) examined the interactive effects of isolation (underlining), readability (understandability), and paragraph organization (intersentence structure) on learning from written instructional materials using a mc test as the criterion. Only paragraph organization had a significant effect on learning; no interactions among the three variables were detected. Carver (1975) assessed the effect of manipulating attention during prose reading using a reading input technique called programmed prose (similar to cloze and programmed text). Programmed prose increased comprehension in low motivation conditions but acted as a distractor in high motivation conditions.

Kintsch et al. (1975) presented short English texts, controlled for number of words and propositions but differing in word concepts in the text base, to college students to assess the function of several content variables influencing comprehension and recall of text material. Number of word concepts affected reading time and recall. Superordinate propositions were recalled better than subordinate propositions and forgotten less when recall was delayed. Recall of word concepts increased as a function of the number of repetitions of the concept in the text base and the number of repetitions of the corresponding words in the actual text. The effects held for listening as well as reading.

Experienced readers report the ability to remember where certain information is located in passages and books. Christie and Just (1976) used ingeniously constructed text passages to examine how subjects remember the location of a sentence in a passage and how they make use of that information in retrieving content information from text. They found location of material could be remembered but did not guarantee memory for content. Also, one does not necessarily have to know where the information is located in a text to retrieve it; although location information may be helpful in some situations. Response latencies were similar for organized as well as disorganized content (Schumacher, Liebert and Fass (1975) reported that disorganized presentation of prose may in some instances facilitate recall). DiVesta and Gray (1973) also demonstrated that scrambled passages were remembered significantly better than a completely logical ordering of the passage. People internalize coherent passages and may also reorganize incoherent passages for internal storage. Eye movement study was found to be a useful technique in researching use of text location in recall.

Yekovich and Kulhavey (1976) reported structural and contextual effects in the organization of prose providing support for the notion of active processing in learning and the interdependency of context and organizational forms. Immediate and delayed word recall and recognition revealed organization form significantly affected retention and memory when subjects read passages organized around either name or attribute sequences. Conceptually organized material facilitated recall; subjects encountering coherently organized material are more likely to learn and remember than when they encounter haphazardly constructed material.

Schallert (1976) investigated the role of context in determining how much a person remembers from prose and exactly what is remembered using ambiguous paragraphs whose semantic representation can be influenced by differing contexts. Task instructions and time of exposure to passages were manipulated to induce differential levels of processing and influence amount of material retained. Recall and recognition tests revealed subjects remembered more information and more context consonant information when given instructions which required processing the paragraphs at a semantic level. Context was thus a powerful factor in determining which meaning was remembered from polysemous paragraphs only when incoming information was processed at a deeper, more semantic level.

References

Adejumo, J. A. The effects of some structured strategies on the acquisition and recall of instructional prose materisl. (Doctoral dissertation, Syracuse University, 1974.) *Dissertation Abstracts International*, 1975, *36*, 180A. (University Microfilms No. 75-13,954.)

Aiken, E. G., Thomas, G. S., & Shennum, W. A. Memory for a lecture: Effects of notes, lecture rate and information density. *Journal of Educational Psychology*, 1975, *67*, 439-444.

Allen, A. L. R. The effect of student contol of course content upon achievement and attitude in a college reading improvement program. (Doctoral dissertation, Ball State University, 1975.) *Dissertation Abstracts International*, 1976, *36*, 5137A. (University Microfilms No. 76-4566.)

Allen, A. R., & Harshbarger, M. Student locus of control and teaching style in relation to college reading improvement. *Reading Improvement*, 1977, *14*, 104-108.

Ambrosino, R. J., Brading, P. L., & Noval, Lorraine K. Reading: A potential source of academic difficulty in medical school. *Journal of Reading Behavior*, 1974, *6*, 367-373.

Anderson, R. C. Control of student mediating processes during verbal learning and instruction. *Review of Educational Research*, 1970, *40*, 349-369.

Anderson, T. H. Cloze measures as indices of achievement comprehension when learning from extended prose. *Journal of Educational Measurement*, 1974, *11*, 83-92.

Anderson, R. C., & Biddle, W. B. On asking people questions about what they are reading. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 9). New York: Academic Press, 1975.

Andrews, L. Reading comprehension and three modes of reading assistance. *Journal of Reading Behavior*, 1975, *5*, 237-241.

Annis, L., & Davis, J. K. The effect of encoding and an external memory device on note taking. *Journal of Experimental Education*, 1975, *44*, 44-46.

Arkes, H. R., Schumacher, G. M., & Gardner, E. T. Effects of orienting tasks on the retention of prose material. *Journal of Educational Psychology*, 1976, *68*, 536-545.

Arlin, M. One-study publishing typifies educational inquiry. *Educational Researcher*, 1977, *6*, 11-15.

Ash, R. A., & Edgell, S. L. A note on the readability of the position analysis questionaire (PAQ). *Journal of Applied Psychology*, 1975, *60*, 765-766.

Baer, F. B. A comparison of the effects of two methods on the reading vocabulary, comprehension, accuracy and rate of selected students at the George Washington University. (Doctoral dissertation, George Washington University, 1974.) *Dissertation Abstracts International*, 1975, *35*, 7668A. (University Microfilms No. 75-12,606.)

Bain, J. W. A study of the effects on college students' reading achievement of two instructional approaches: Audio-visual and visual. (Doctoral dissertation, American University, 1974.) *Dissertation Abstracts International*, 1974, *35*, 2114A. (University Microfilms No. 74-21,687.)

Ball, F., Wood, C., & Smith, E. E. When are semantic targets detected faster than visual or acoustic ones? *Perception and Psychophysics*, 1975, *7*, 1-8.

Baron, J., & McKillop, B. J. Individual differences in speed of phonemic analysis, visual analysis and reading. *Acta Psychologica*, 1975, *39*, 91-96.

Bankston, L. V. C. The effects of the readability of mathematical materials on achievement in remedial mathematics in a selected community college. (Doctoral dissertation, University of Texas at Austin, 1975.) *Dissertation Abstracts International*, 1975, *36*, 815A. (University Microfilms No. 75-16,633.)

Bassin, C. B. Telegraphic prose: The effects of three deletion schemes on comprehension and reading rate of three types of literature. (Doctoral dissertation, Texas A. & M. University, 1975.) *Dissertation Abstracts International*, 1976, *36*, 5139A. (University No. 75-25,081.)

Bassin, C. B., & Martin, C. J. Effect of three types of redundancy reduction on comprehension, reading rate, and reading time of English prose. *Journal of Educational Psychology*, 1976, *68*, 649-652.

Beneke, W. N., & Harris, M. B. Teaching self-control of study behaviors. *Behavior Research and Therapy*, 1972, *10*, 34-41.

Blanton, W. E., & Smith, C. F., Jr. 1974 Review of research on college-adult reading. *Reading: Convention and inquiry.* National Reading Conference Yearbook, 1975, *24*, 1-66.

Blumenfeld, W. S., Black, W. R., & Morris, S. L. Interrater reliability of the Gunning readability index: Methodological note. *Perceptual and Motor Skills*, 1975, *40*, 902.

Blumenfeld, W. S., & Justin, B. M. Six replicated investigations of the relationship between Flesch and Gunning readability indices. *Perceptual and Motor Skills*, 1975, *40*, 110.

Boker, J. R. Immediate and delayed retention effects of interspersing questions in written instructional passages. *Journal of Educational Psychology*, 1974, *66*, 96-98.

Booher, H. R. Relative comprehensibility of pictorial information and printed words in proceduralized instructions. *Human Factors*, 1975, *17*, 266-277.

Botwinick, J., & Storandt, M. Vocabulary ability in later life. *Journal of Genetic Psychology*, 1974, *125*, 303-308.

Bostian, L. R. Effect of line width on reading speed and comprehension. *Journalism Quarterly*, 1976, *53*, 328-330.

Bower, G. H. Experiments on story understanding and recall. *Quarterly Journal of Experimental Psychology*, 1976, *28*, 511-534.

Bradshaw, J. L. Three interrelated problems in reading: A review. *Memory and Cognition*, 1975, *3*, 123-134.

Brandt, J. D. Internal versus external locus of control and performance in controlled and motivated reading-rate improvement instructions. *Journal of Counseling Psychology*, 1975, *22*, 377-383.

Brown, L. A. The effect of isolation, readability and paragraph organization on learning from written instructional materials. (Doctoral dissertation, University of Nebraska, Lincoln, 1974.) *Dissertation Abstracts International*, 1975, *36*, 772A-773A. (University Microfilms No. 75-16,884.)

Bryant, B. E., Currier, F. P., & Morrison, A. J. Relating life style factors of person to his choice of a newspaper. *Journalism Quarterly*, 1976, *53*, 74-79.

Bull, S. G. The role of questions in maintaining attention to textual material. *Review of Educational Research*, 1973, *43*, 83-87.

Bull, S. G., & Disney, H. F. Epistemic-curiosity-arousing prequestions: Their effect on long term retention. *Journal of Educational Psychology*, 1973, *65*, 45-49.

Burgess, B. A., Cranney, A. G., & Larsen, J. J. Effect on academic achievement of a voluntary university reading program. *Journal of Reading*, 1976, *19*, 644-646.

Carpenter, P. A. On the comprehension, storage, and retriieval of comparative sentences. *Journal of Verbal Learning and Verbal Behavior*, 1974, *13*, 401-411.

Carpenter, P. A., & Just, M. A. Sentence comprehension: A psycholoinguistic processing model of verification. *Psychological Review*, 1975, *82*, 45-73.

Carpenter, P. A., & Just, M. A. Models of sentence verification and linguistic comprehension. *Psychological Review*, 1976, *83*, 318-322.

Carpenter, T. W., & Jones, Y. Improving comprehension rate gain at the college level. *Journal of Reading*, 1975, *19*, 223-225.

Carroll, J. B. *Learning from verbal discourse in educational media: A review of the literature.* Princeton, N. J.: Educational Testing Service, 1971.

Carter, J. F., & Van Matre, N. H. Note taking versus note having. *Journal of Educational Psychology*, 1975, *67*, 900-904.

Carver, R. P. Measuring the primary effect of reading: Reading storage technique, understanding judgments, and cloze. *Journal of Reading Behavior*, 1974, *6*, 249-274.

Carver, R. P. Comparing the reading-storage test to the paraphrase test as measures of the primary effect of prose reading. *Journal of Educational Psychology*, 1975, *67*, 274-284.

Carver, R. P. Manipulation of attention during reading using programmed prose. *Journal of REading Behavior*, 1975, *7*, 241-254.

Carver, R. P. Measuring prose difficulty using the Rauding scale. *Reading Research Quarterly*, 1975-1976, *11*, 660-683. (a)

Carver, R. P. A response to Kibby's criticisms of the Rauding scale research. *Reading Research Quarterly*, 1975-1976, *11*, 706-715. (b)

Carver, R. P. Word length, prose difficulty, and reading rate. *Journal of Reading Behavior*, 1976, *8*, 193-203.

Carstens, R. W., & McKeag, R. A. Cloze test scores for assessing reading comprehension: Some considerations for interpreting data. *Reading Improvement*, 1975, *12*, 8-10.

Catlin, J., & Jones, N. K. Verifying affirmative and negative sentences. *Psychological Review*, 1976, *83*, 497-501.

Christie, J. M., & Just, M. A. Remembering the location and content of sentences in a prose passage. *Journal of Educational Psychology*, 1976, *68*, 702-710.

Cochran, J. A. A descriptive study examining grades, reading ability, and attrition of certain low-income freshmen at Arizona State University. (Doctoral dissertation, Arizona State University, 1974.) *Dissertation Abstracts International*, 1975, *35*, 4042A-4043A. (University Microfilms No. 74-20,127.)

Cochran, M. L. The revised Ohio literacy test with normal adults. *Journal of Clinical Psychology*, 1974, *30*, 321-325.

Coke, E. U. The effects of readability on oral and silent reading rate. *Journal of Educational Psychology*, 1974, *66*, 406-409.

Coleman, M., & Liau, T. L. A computer readability formula designed for machine scoring. *Journal of Applied Psychology*, 1975, *60*, 283-284.

Cook, V. J. Strategies in the comprehension of relative clauses. *Language and Speech*, 1975, *18*, 204-212.

Corlett, D. Library skills, study habits and attitudes, and sex as related to academic achievement. *Educational and Psychological Measurement*, 1974, *34*, 967-969.

Cosky, J. J. The role of letter recognition in word recognition. *Memory and Cognition*, 1976, *4*, 207-214.

Cox, J. S. R. A study of the syntactic competence of adult beginning readers. (Doctoral dissertation, University of North Carolina at Chapel Hill, 1975.) *Dissertation Abstracts International*, 1976, *36*, 6525A. (University Microfilms No. 76-9230.)

Coy, J. R. J. Initial validation of the adult learning rate test. (Doctoral dissertation, University of Northern Colorado, 1975.) *Dissertation Abstracts International*, 1975, *36*, 663A-664A. (University Microfilms No. 75-16,835.)

Cronbach, L. J., & Snow, R. E. *Aptitudes and instructional methods*. New York: Irvington Publishers, Inc., 1977.

Crouse, J. H. Acquisition of college course material under conditions of repeated testing. *Journal of Educational Psychology*, 1974, *66*, 367-372.

Dallas, M., & Merikle, P. M. Response processes and semantic-context effects. *Bulleting of the Psychonomic Society*, 1976, *8*, 441-444.

Davies, G., & Proctor, J. The recall of concrete and abstract sentences as a function of interpolated task. *British Journal of Psychology*, 1976, *67*, 63-72.

Denne, K. T. Reading proficiency: A new dilemma in medical college admissions. (Doctoral dissertation, University of Utah, 1974.) *Dissertation Abstracts International*, 1975, *35*, 4064A. (University Microfilms No. 74-29,548.)

DePietro, A.H. An investigation of interaction analysis and selected school and student inputs toward the prediction of literacy achievement, attendance, and dropout incidence in adult basic education projects. (Doctoral dissertation, Catholic University of America, 1975.) *Dissertation Abstracts International*, 1975, *36*, 1241A. (University Microfilms No. 75-18,530.)

DiVesta, F. J., & Gray, G. S. Listening and note-taking: II. Immediate and delayed recall as functions of variations in thematic continuity, note taking, and length of listening-review intervals. *Journal of Educational Psychology*, 1973, *64*, 278-287.

Drummond, R. J., Smith, R. K., & Pinette, C. A. Internal-external control construct and performance in an individualized community college reading course. *Reading Improvement*, 1975, *12*, 34-38.

Duchastel, P. C., & Brown, B. R. Incidental and relevant learning with instructional objectives. *Journal of Educational Psychology*, 1974, *66*, 481-485.

Duchastel, P. C., & Merrill, P. F. The effects of behavioral objectives upon learning: A review of empirical studies. *Review of Educational Research*, 1973, *43*, 53-69.

Duell, O. K. Effect of types of objective, level of test questions, and judged importance of test materials upon posttest performance. *Journal of Educational Psychology*, 1975, *67*, 380-384.

Dulin, K. L., & Greenwald, M. J. Mature readers' affective response to three specific propaganda devices: Loaded words, name-calling, and borrowed prestige/borrowed dislike. *Reading Convention and inquiry*. National Reading Conference Yearbook, 1975, *24*, 267-272.

Dyer, J. W. Effects of notetaking and repetition on encoding and retention of printed matter. (Doctoral dissertation, Arizona State University, 1974.) *Dissertation Abstracts International*, 1975, *35*, 4246A. (University Microfilms No. 75-1665.)

Edfeldt, A. W. Concerning differences between good and poor readers. *Journal of Experimental Education*, 1975, *43*, 90-93.

Edwards, R. D. The cloze precedure as a measure of the reading comprehension of poetry. (Doctoral dissertation, University of Arizona, 1973.) *Dissertation Abstracts International*, 1974, *34*, 4982A. (University Microfilms No. 74-4182.)

Eriksen, B. A., & Eriksen, C. W. The importance of being first: A tachistoscopic study of the contribution of each letter to the recognition of four-letter words. *Perception and Psychophysics*, 1974, *15*, 66-72.

Fairbanks, M. M. The effect of college reading improvement programs on academic achievement. *Interaction: Research and practice in college-adult reading*. National Reading Conference Yearbook, 1974, *23*, 105-114.

Fairbanks, M. M. Relationship between research control and reported results of college reading improvement programs. *Reading: Convention and inquiry*. National Reading Conference Yearbook, 1975, *24*, 80-93.

Farley, F. H., & Yen, S. Arousal and cognition: Word arousal and visual search. *Perceptual and Motor Skills*, 1976, *43*, 699-702.

Farrell, T. J. Reading in the community college. *College English*, 1975, *37*, 40-46.

Faw, H. W., & Waller, T. G. Mathemagenic behaviors and efficiency in learning from prose materials: Review, critique, and recommendations. *Review of Educational Research*, 1976, *46*, 691-720.

Felker, D. B., & Dapra, R. A. Effects of questions type and question placement on problem-solving ability from prose material. *Journal of Educational Psychology*, 1975, *67*, 380-384.

Finn, P. J. Deriving questions as a means of assessing the amount of information in a written passage. *Interaction: Research and practice in college-adult reading*. National Reading Conference Yearbook, 1974, *23*, 205-208.

Fisher, J. L., & Harris, M. B. Effect of note taking and review on recall. *Journal of Educational Psychology*, 1973, *65*, 321-325.

Fisher, J. L., & Harris, M. B. Effect of note-taking preference and type of notes taken on memory. *Psychological Reports*, 1974, *35*, 384-386.

Fisher, J. L., & Harris, M. B. Note taking and recall. *Journal of Educational Research*, 1974, *67*, 291-292.

Fleming, J. T. Alternative interpretations of evidence for phonemic recoding in visual word recognition. *Journal of Reading Behavior*, 1976, *8*, 7-18.

Florio, C. B. An assessment of the effectiveness of the remedial reading courses at San Antonio College. (Doctoral dissertation, University of Texas at Austin, 1975.) *Dissertation Abstracts International*, 1975, *36*, 2664A. (University Microfilms No. 75-24,864.)

Ford, B. C. Factors involved in the reading ability of students in Ontario colleges of applies arts and technology. (Doctoral dissertation, University of Toronto (Canada(, 1972.) *Dissertation Abstracts International*, 1974, *34*, 4694A. (To obtain a microfilm copy please order directly FROM THE National Library of Canada at Ottawa.)

Forster, K. I., & Bednall, E. S. Terminating and exhaustive search in lexical access. *Memory and Cognition*, 1976, *4*, 53-61.

Frase, L. T. Effect of question location, pacing and mode upon retention of prose material. *Journal of Educational Psychology*, 1968, *59*, 244-249.

Frase, L. T. Boundary conditions for mathemagenic behaviors. *Review of Educational Research*, 1970, *40*, 337-348.

Frase, L. T. Maintenance and control in the acquisition of knowledge from written materials. In R. O. Freedle and J. B. Carroll (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D.C.: V. H. Winston and Sons, 1972.

Frase, L. T. Sampling and response requirements of adjunct questions. *Journal of Educational Psychology*, 1973, *65*, 273-278.

Frase, L. T. On learning from prose. *Educational Psychologist*, 1973, *10*, 15-23.

Frase, L. T. Prose processing. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 9). New York: Academic Press, 1975.

Frase, L. T., & Kreitzberg, V. S. Effects of topical and indirect learning directions on prose recall. *Journal of Educational Psychology*, 1974, *67*, 320-324.

Frauenheim, J. G. A follow-up study of adult males who were clinically diagnosed as dyslexic in childhood. (Doctoral dissertation, Wayne State University, 1975.) *Dissertation Abstracts International*, 1975, *36*, 2741A. (University Microfilms No. 75-25,240.)

Friedman, M. P., & Wilson, R. W. Application of unobtrusive measures to the study of textbook usage by college students. *Journal of Applied Psychology*, 1975, *60*, 659-662.

Fronk, R. H. The effect of student reading level, type of instructional material, and mode of instruction on time required to reach mastery level learning. (Doctoral dissertation, University of Iowa, 1974.) *Dissertation Abstracts International*, 1974, *35*, 2073A. (University Microfilms No. 74-21,890.)

Gagne, E. D., & Rothkopf, E. Z. Text organization and learning goals. *Journal of Educational Psychology*, 1975, *67*, 445-450.

Geis, L. K. Development and standardization of parallel forms of a syllabication skills test. (Doctoral dissertation, University of Oklahoma, 1975.) *Dissertation Abstracts International*, 1976, *36*, 6549A. (University Microfilms No. 76-8330.)

Gilbert, T. The effect of dictated notes on the recall and comprehension of connected and disconnected verbal material. *British Journal of Educational Technology*, 1975, *6*, 62-65.

Glass, G. V. Primary, secondary and meta-analysis of research. *Educational Researcher*, 1976, *5*, 3-8.

Gnewuch, M. M. The effect of vocabulary training upon the development of vocabulary, comprehension, total reading, and rate of reading of college students. (Doctoral dissertation, Oklahoma State University, 1973.) *Dissertation Abstracts International*, 1974, *34*, 6254A. (University Microfilms No. 74-8019.)

Golinkoff, R. M. A comparison of reading comprehensio processes in good and poor comprehenders. *Reading Research Quarterly*, 1975-1976, *11*, 622-659.

Golinkoff, R. M., & Rosinski, R. R. Decoding, semantic processing and reading comprehension skill. *Child Development*, 1976, *47*, 252-258.

Green, D. W., & Shallice, T. Direct visual access in reading for meaning. *Memory and Cognition*, 1976, *4*, 753-758.

Gurrola, S. Determination of the relative effectiveness and efficiency of selected combinations of SQ3R study method components. (Doctoral dissertation, New Mexico State University, 1974.) *Dissertation Abstracts International*, 1975, *35*, 6938A. (University Microfilms No. 75-10,822.)

Hakes, D. T., Evens, J. S., & Brannon, L. L. Understanding sentences with relative clauses. *Memory and Cognition*, 1976, *4*, 283-290.

Hakstian, A. R., & Cattell, R. B. The checking of primary ability structure on a broader basis of performance. *British Journal of Educational Psychology*, 1974, *44*, 140-154.

Hansen, D. M. A discourse structure analysis of the comprehension of rapid readers. (Doctoral dissertation, Brigham Young University, 1975.) *Dissertation Abstracts International*, 1976, *36*, 7293A. (University Microfilms No. 76-9839.)

Harris, R. J. Comprehension of metaphors: A test of the two-stage processing model. *Bulletin of the Psychonomic Society*, 1976, *8*, 312-314.

Haselton, S. S. Identification of reading and study skills related to academic success in each of the school's at Oregon State University. (Doctoral dissertation, Oregon State University, 1974.) *Dissertation Abstracts International*, 1974, *35*, 2007A. (University Microfilms No. 74-23,440.)

Haskins, J. B. Pretesting interest in messages. *Journal of Advertising Research*, 1975, *15*, 31-35.

Haviland, M. R. Provision for adult literacy instruction in England. (Doctoral dissertation, University of Massachusetts, 1973.) *Dissertation Abstracts International*, 1974, *34*, 6257A. (University Microfilms No. 74-8598.)

Henard, K. F. Life change and reading achievement as predictors of academic performance for selected community college freshmen. (Doctoral dissertation, Texas A. & M. University, 1975.) *Dissertation Abstracts International*, 1976, *36*, 5082A. (University Microfilms No. 76-3644.)

Henderson, M. A. Prescriptive and personalized teaching methods in community and junior college reading courses. (Doctoral dissertation, University of Oklahoma, 1974.) *Dissertation Abstracts International*, 1975, *35*, 5778A. (University Microfilms No. 75-6524.)

Henderson, M. A. Individualized reading instruction in two-year colleges. *Journal of Reading*, 1976, *19*, 464-471.

Hennessy, J. J., & Loveless, E. The contribution of three reading tests to a community college placement program. *Educational and Psychological Measurement*, 1976, *36*, 459-464.

Herndon, M. A. An approach toward computer control of redundancy in textual materials. *Journal of Reading Behavior*, 1976, *8*, 259-271.

Hesson, J. C. Effects of performance objectives on reading achievement in a community college reading program. (Doctoral dissertation, University of New Mexico, 1974.) *Dissertation Abstracts International*, 1975, *35*, 5681A. (University Microfilms No. 75-6967.)

Hiller, J. H. Learning from prose text: effects of readability level, inserted question difficulty, and individual differences. *Journal of educational Psychology*, 1974, *66*, 202-211.

Hiscock, M. Effects of adjective imagery on recall from prose. *Journal of General Psychology*, 1976, *94*, 295-299.

Hoiberg, A., Hysham, c. J., & Berry, N. H. Effectiveness of recruits assigned to academic remedial training. *Psychological Reports*, 1974, *35*, 1007-1014. (a)

Hoiberg, A., Hysham, C. J., & Berry, N. H. The neuropsychiatric implications of illiteracy–twenty years later. *Journal of Clinical Psychology*, 1974, *30*, 533-535. (b)

Holyoak, K. J. The role of imagery in the evaluation of sentences: imagery or semantic factors? *Journal of Verbal Learning and Verbal Behavior*, 1974, *13*, 163-166.

Hornsby, J. G. The relationship between verbal responses, reading achievement, intelligence, and performance ability for academic and vocational junior college students. (Doctoral dissertation, University of Southern Mississippi, 1973.) *Dissertation Abstracts International*, 1974, *34*, 4563A. (University Microfilms No. 74-3935.)

Howard, V. P. H. Developments in instruction in selected four year college reading improvement programs throughout the U.S., 1950-74. (Doctoral dissertation, George Washington University, 1975.) *Dissertation Abstracts International*, 1975, *36*, 2667A. (University Microfilms No. 75-25,387.)

Howe, M. J. A. The utility of taking notes as an aid to learning. *Educational Research*, 1974, *16*, 222-227.

Howe, M. J. A., & Colley, L. Retroactive interference in meaningful learning. *British Journal of Educational Psychology*, 1976; *46*, 26-30. (a)

Howe, M. J. A., & Colley, L. The influence of questions encountered earlier on learning from prose. *British Journal of Educational Psychology*, 1976, *46*, 149-154. (b)

Howe, M. J. A., & Singer, L. Presentation variables and students' activities in meaningful learning. *British Journal of Educational Psychology*, 1975, *45*, 52-61.

Hultsch, D. F., & Craig, E. R. Adult age differences in the inhibition of recall as a function of retrieval cues. *Developmental Psychology*, 1976, *12*, 83-84.

Hutchison, L. F. E. The effects of instructional strategies related to preferred interests of functionally illiterate adults. (Doctoral dissertation, Ball State University, 1974.) *Dissertation Abstracts International*, 1974, *35*, 1946A-1947A. (University Microfilms No. 74-22,803.)

Idstein, P., & Jenkins, J. R. Underlining vs. repetitiv reading. *Journal of Educational Research*, 1972, *65*, 321-323.

Jackson, M. D., & McClelland, J. L. Sensory and cognitive determinants of reading speed. *Journal of Verbal Learning and Verbal Behavior*, 1975, *14*, 565-574.

Jenkins, J. R., & Bausell, R. B. Cognitive structure variables in prose learning. *Journal of Reading Behavior*, 1976, *8*, 47-66.

Jenkins, J. R., & Neisworth, J. T. The facilitative influence of instructional objectives. *Journal of Educational Research*, 1973, *66*, 254-256.

Johnson, B. F. The development and evaluation of a model for computer-assisted reading instruction in adult basic education programs. (Doctoral dissertation, Florida State University, 1974.) *Dissertation Abstracts International*, 1975, *35*, 4131A. (University Microfilms No. 75-960.)

Johnson, H. C. The effects of subjective organizational ability and the organization of materials on reading comprehension: an information processing interpretation. (Doctoral dissertation, Michigan State University, 1974.) *Dissertation Abstracts International*, 1975, *35*, 5925A. (University Microfilms No. 75-7190.)

Johnson, R. E. Abstractive processes in the remembering of prose. *Journal of Educational Psychology*, 1974, *66*, 772-779.

Just, M. A., & Carpenter, P. A. The relation between comprehending and remembering some complex sentences. *Memory and Cognition*, 1976, *4*, 318-322.

Kahn, G. B. The development and implementation of an applied psychology course utilizing behavior modification procedures to supplement a community college reading program for open enrollment students. (Doctoral dissertation, Columbia University Teachers College, 1975.) *Dissertation Abstracts International*, 1976, *36*, 6446A-6447A. (University Microfilms No. 76-7777.)

Kamm, K., Otto, W., & Dulin, K. L. An evaluation of reading improvement workbooks. *Interaction: research and practice in college-adult reading.* National Reading Conference Yearbook, 1974, *23*, 243-248.

Kammann, R. The comprehensibility of printed instructions and the flow-chart alternative. *Human Factors*, 1975, *17*, 183-191.

Kaplan, R. Effects of learning with part vs. whole presentations of instructional objectives. *Journal of Educational Psychology*, 1974, *66*, 787-792.

Kaplan, R., & Rothkopf, E. Z. Instructional objectives as directions to learners: effect of passage length and amount of objective relevant content. *Journal of Educational Psychology*, 1974, *66*, 448-456.

Kaplan, R., & Simmons, F. G. Effects of instructional objectives used as orienting stimuli or as summary review upon prose learning. *Journal of Educational Psychology*, 1974, *66*, 614-622.

Karlin, M. B., & Bower, G. H. Semantic category effects in visual word search. *Perception and Psychophysics*, 1976, *19*, 417-424.

Keetz, M. A. Ability of students to identify correct responses before reading: a replication. *Reading: convention and inquiry.* National Reading Conference Yearbook, 1975, *24*, 235-240.

Khan, S. B., & Roberts, D. M. Structure of academic attitudes and study habits. *Educational and Psychological Measurement*, 1975, *35*, 835-842.

Kibby, M. W. The proper study of readability: a reaction to Carver's "Measuring prose difficulty using the Rauding scale." *Reading Research Quarterly*, 1975-1976, *11*, 686-705.

King, D. J. Total presentation time and total learning time in connected discourse learning. *Journal of Experimental Psychology*, 1974, *103*, 586-589.

Kinnamon, S. E. Commercial television and adult reading. *Journal of Reading*, 1975, *18*, 470-474.

Kintsch, W., Kozminsky, E., Streby, W. J., McKoon, G., & Keenan, J. M. Comprehension and recall of text as a function of content variable. *Journal of Verbal Learning and Verbal Behavior*, 1975, *14*, 196-214.

Kistulentz, A. C. An analysis of two approaches used in teaching reading and study skills improvement in a two-year community college. (Doctoral dissertation, Pennsylvania State University, 1975.) *Dissertation Abstracts International*, 1976, *36*, 4227A-4228A. (University Microfilms No. 76-1360.)

Klare, G. R. A second look at the validity of readability formulas. *Journal of Reading Behavior*, 1976, *8*, 129-152.

Kleiman, G. M. Speech recoding in reading. *Journal of Verbal Learning and Verbal Behavior*, 1975, *14*, 323-339.

Knowles, J. H. Similarities and differences in characteristics of students participating in two types of adult literacy programs. (Doctoral dissertation, Florida State University, 1974.) *Dissertation Abstracts International*, 1975, *35*, 4115A. (University Microfilms No. 75-936.)

Kolers, P. A. Specificity in operations in sentence recognition. *Cognitive Psychology*, 1975, *7*, 289-306.

Koran, M. L., & Koran, J. J., Jr. Interaction of learner aptitudes with question pacing in learning from prose. *Journal of Educational Psychology*, 1975, *67*, 76-82.

Krulee, G. K., & Ramsburg, R. E. Semantic anomalies and sentence recognition. *Perceptual and Motor Skills*, 1974, *38*, 1275-1286.

Kurzman, M. The reading ability of college freshmen compared to the readability of their textbooks. *Reading Improvement*, 1974, *11*, 13-25.

Ladas, H. The mathemagenic effects of factual review questions on the learning of incidental information. *Review of Educational Research*, 1973, *43*, 71-82.

Lamb, C. W. An investigation of growth in critical reading ability of adult basic education students. (Doctoral dissertation, Florida State University, 1975.) *Dissertation Abstracts International*, 1975, *36*, 3334A. (University Microfilms No. 75-26,790.)

LaPorte, R. E., & Nath, R. Role of performance goals in prose learning. *Journal of Educational Psychology*, 1976, *68*, 260-264.

Larkin, J. H., & Reif, F. Analysis and teaching of a general skill for studying scientific text. *Journal of Educational Psychology*, 1976, *68*, 431-440.

Larsen, J. J., Tillman, C. E., & Cranney, A. G. Trends in college freshman reading ability. *Journal of Reading*, 1976, *19*, 367-369.

Lawson, T. L. Effects of instructional objectives on learning and retention. *Instructional Science* 1974, *3*, 1-21.

Layton, P., & Simpson, A. J. Surface and deep structure in sentence comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1975, *14*, 658-664.

Lazerson, B. H. The influence of highly variable spelling upon the reading performance of skilled readers of Modern English. *Reading Research Quarterly*, 1974-1975, *10*, 583-615.

Levy, B. A. Vocalization and suppression effects in sentence memory. *Journal of Verbal Learning and Verbal Behavior*, 1975, *14*, 304-316.

Liau, T. L., Bassin, C. B., Martin, C. J., & Coleman, E. B. Modifications of the Coleman readability formulas. *Journal of Reading Behavior*, 1976, *8*, 381-386.

Littrell, J. H. Reducing phonic disability in junior college. *Journal of Reading*, 1976, *20*, 150-152.

Livingston, C. L. A study of the effects of different levels of initial diagnosis on gains in reading achievement for adults in a community college. (Doctoral dissertation, University of Illinois at Urbana-Champaign, 1975.) *Dissertation Abstracts International*, 1975, *36*, 2474A. (University Microfilms No. 75-24,357.)

Lowry, D. T., & Marr, T. J. Clozentropy as a measure of international communication comprehension. *Public Opinion Quarterly*, 1975, *39*, 301-312.

Lund, P. A., & Ivanoff, J. M. Correspondence of self-concept measures with levels of reading achievement. *Journal of Reading Behavior*, 1974, *6*, 159-165.

McConkie, G. W., & Meyer, B. J. F. Investigation of reading strategies: II. A replication of payoff condition effects. *Journal of Reading Behavior*, 1974, *6*, 151-158.

McCreath, E. E. An investigation of reading habits, reading interests, and their relationship to reading improvement in an urban opendoor junior college. *Reading: Convention and inquiry*. National Reading Conference Yearbook, 1975, *24*, 100-105.

McNinch, G., Kazelskis, R., & Cox, J. A. Appropriate cloze deletion schemes for determining suitability of college textbooks. *Interaction: Research and practice in college-adult reading*. National Reading Conference Yearbook, 1974, *23*, 249-253.

McWhorter, K. T. The influence of passage organizational structure upon two estimates of readability. (Doctoral dissertation, State University of New York at Buffalo, 1974.) *Dissertation Abstracts Intrnational*, 1974, *35*, 3303A-3304A. (University Microfilms No. 74-29,242.)

Meader, J. Effects of environmental controls on skill development and transfer in prose learning. (Doctoral dissertation, University of California, Los Angeles, 1975.) *Dissertation Abstracts International*, 1975, *36*, 685A-686A. (University Microfilms No. 75-16,790.)

Mahoney, M. M. A comparison of three methods of teaching reading-study skills at the college level. (Doctoral dissertation, Boston University School of Education, 1974.) *Dissertation Abstracts International*, 1974, *35*, 847A. (University Microfilms No. 74-17,927.)

Manikas, W. T. A remedial program for underprepared, black junior college students. (Doctoral dissertation, Florida Atlantic University, 1974.) *Dissertation Abstracts International*, 1974, *35*, 2011A. (University Microfilms No. 74-22,195.)

Marcel, T. The effective visual field and the use of context in fast and slow readers of two ages. *British Journal of Psychology*, 1974, *65*, 479-492.

Mason, J. M. The roles of orthographic, phonological and word frequency variables on word-nonword decisions. *American Educational Research Journal*, 1976, *13*(3), 199-206.

Mavrogenes, N. A., Winkley, C. K., Hanson, E., & Vaaca, R. T. Concise guide to standardized secondary and college reading tests. *Journal of Reading*, 1974, *18*, 12-22.

Maxwell, M. J. Developing a learning center: Plans, problems and progress. *Journal of Reading*, 1975, *18*, 462-469.

Mayer, R. E. Forward transfer of different reading strategies evoked by testlike events in mathematics text. *Journal of Educational Psychology*, 1975, *67*, 165-169.

Meyer, D. E., Schvaneveld, R. W., & Ruddy, M. C. Functions of graphemic and phonemic codes in visual word recognition. *Memory and Cognition*, 1974, *2*, 309-321.

Miller, L. R. Predictive powers of the Flesch and Bormuth readability formulas. *Journalism Quarterly*, 1974, *51*, 508-511.

Millott, R., & Cranney, A. G. Personality correlates of college reading and study skills. *Journal of Reading Behavior*, 1976, *8*, 335-336.

Mistler-Lachman, J. L. Depth of comprehension and sentence memory. *Journal of Verbal Learning and Verbal Behavior*, 1974, *13*, 98-106.

Mizer, M. M. Differences between educationally active and noneducationally active older adults in reading, interpersonal values, and life satisfaction. (Doctoral dissertation, Arizona State University, 1975.) *Dissertation Abstracts International*, 1975, *36*, 2507A-2508A. (University Microfilms No. 75-24,225.)

Morrow, C. A. The effects of selected transformations on the comprehension of college-level expository prose. (Doctoral dissertation, University of Akron, 1974.) *Dissertation Abstracts International*, 1974, *35*, 2013A-2014A. (University Microfilms No. 74-21,086.)

Morse, J. M. Effect of reader-generated questions on learning from prose. (Doctoral dissertation, Rutgers University, The State University of New Jersey, 1975.) *Dissertation Abstracts International*, 1976, *36*, 6635A-6636A. (University Microfilms No. 76-8701.)

Muma, J. R., Perry, M. E. A., & Gallagher, J. W. Some semantic properties in sentence perception. *Psychological Reports*, 1974, *35*, 23-32.

Musik, S. K. The relation of the self-concept to reading. (Doctoral dissertation, U.S. International University, 1974.) *Dissertation Abstracts International*, 1974, *35*, 324A. (University Microfilms No. 74-14,318.)

Newstead, S. E. Semantic constraints and sentence perception. *British Journal of Psychology*, 1976, *67*, 165-173.

Nilagupta, S. The relationship between reading ability in English of Thai college students and the readability levels of their English course textbooks. (Doctoral dissertation, Florida State University, 1975.) *Dissertation Abstracts International*, 1975, *36*, 2077A. (University Microfilms No. 75-21,426.)

O'Donnell, J. F. An experimental study of the effects of the supplemental use of a psycholinguistic remedial tutorial program on the reading and writing behaviors of black high-risk college freshmen and on their attitudes toward reading, writing and other college-related stimuli. (Doctoral dissertation, Temple University, 1974.) *Dissertation Abstracts International*, 1974, *35*, 1552A-1553A. (University Microfilms No. 74-19,763.)

Ohnmacht, F. W., & Fleming, J. T. Further effects of selected deletion strategies and varying contextual constraints on cloze performance. *Interaction: Research and practice in college-adult reading*. National Reading Conference Yearbook, 1974, *23*, 163-171.

Ohnmacht, F. W., & Weiss, F. Construct validity and reading research: Some comments with a data simulation exemplar. *Journal of Reading Behavior*, 1976, *8*, 19-25.

Pace, A. J., & Golinkoff, R. M. Relationship between difficulty and access of single word meaning by skilled and less skilled readers. *Journal of Educational Psychology*, 1976, *68*, 760-767.

Packwood, G. E. The effects of adjunct questions on the retention of topically related and spatially contiguous incidental information. (Doctoral dissertation, University of Missouri-Columbia, 1975.) *Dissertation Abstracts International*, 1976, *36*, 6562A-6563A. (University Microfilms No. 76-7533.)

Palmatier, R. A., & Bennett, J. M. Notetaking habits of college students. *Journal of Reading*, 1974, *18*, 215-218.

Papandrea, J. M. The effectiveness of a remedial program on the academic low-achiever in a community college. (Doctoral dissertation, Northern Illinois University, 1974.) *Dissertation Abstracts International*, 1975, *35*, 4993A. (University Microfilms No. 75-2315.)

Parham, J. A. W. Academic success of persistent and nonpersistent students voluntarily enrolled in a university reading program. *Journal of Reading*, 1977, *20*, 693-696.

Patterson, O. The effect of intersentence factors upon the cloze scores of passages. (Doctoral dissertation, Hofstra University, 1973.) *Dissertation Abstracts International*, 1974, *35*, 266A-267A. (University Microfilms No. 74-16,424.)

Pearson, P. D., & Kamil, M. L. Word recognition latencies as a function of form class, stem length, and affix length. *Visible Language*, 1974, *8*, 241-246.

Pedrini, D. T., & Pedrini, B. C. Reading abilities and college grades. *College Student Journal*, 1975, *9*, 37-42. (a)

Pedrini, D. T., & Pedrini, B. C. College grades and reading abilities. *Reading Improvement*, 1975, *12*, 75-80. (b)

Penrose, J., Weaver, D. H., Cole, R. R., & Shaw, D. L. The newspaper nonreader 10 years later: A partial replication of Westley-Severin. *Journalism Quarterly*, 1974, *51*, 631-638.

Perfetti, C. A., & Hogaboam, T. The relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology*, 1975, *67*, 461-469.

Peters, D. L. Effects of note taking and rate of presentation on short-term objective test performance. *Journal of Educational Psychology*, 1972, *63*, 276-280.

Pezdek, K., & Royer, J. M. The role of comprehension in learning concrete and abstract sentences. *Journal of Verbal Learning and Verbal Behavior*, 1974, *13*, 551-558.

Phlegar, B. L. L. An attitudinal study of teachers and students in content area reading in the community college. (Doctoral dissertation, Virginia Polytechnic Institute and State University, 1974.) *Dissertation Abstracts International*, 1975, *35*, 7689A. (University Microfilms No. 75-11,947.)

Pincus, S., & Waters, L. K. Effect of age restrictions and pornographic content on desirability of reading material. *Psychological Reports*, 1976, *38*, 943-947.

Postman, L. Experimental analysis of learning to learn. In G. H. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation*. New York: Academic Press, 1969.

Pyrczak, F. Passage-dependence of reading comprehension questions: Examples. *Journal of Reading*, 1975, *18*, 308-311.

Pyrczak, F. Readability for "Instructions for Form 1040." *Journal of Reading*, 1976, *20*, 121-127.

Rakes, T. A. A comparative readability study of materials used to teach adults to read. *Adult Education*, 1973, *3*, 192-202.

Rakes, T. A., Lintz, E., & Payton, L. An evaluation of the reading achievement of disadvantaged freshmen enrolled in competency-based and traditional college reading improvement classes. *Interaction: Research and practice in college-adult reading*. National Reading Conference Yearbook, 1974, *23*, 280-283.

Ramig, C. J., & Dulin, K. L. A one-year reading methods seminar for adult basic education teachers. *Reading: Convention and inquiry*. National Reading Conference Yearbook, 1975, *24*, 137-142.

Rayner, K. The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 1975, *7*, 65-81.

Rayner, K. Developmental changes in word recognition strategies. *Journal of Educational Psychology*, 1976, *68*, 323-329.

Reichurdt, K. W. Playing dead or running away--defense reactions during reading. *Journal of Reading*, 1977, *20*, 706-711.

Reynolds, A. G., & Flagg, P. W. Recognition memory for elements of sentences. *Memory and Cognition*, 1976, *4*, 422-432.

Richards, C. S. Behavior modification of studying through study skills advice and self-control procedures. *Journal of Counseling Psychology*, 1975, *5*, 431-436.

Richardson, J. T. E. The effects of stimulus attributes upon latency of word recognition. *British Journal of Psychology*, 1976, *67*, 315-325.

Rickards, J. F. Processing effects of advance organizers interspersed in text. *Reading Research Quarterly*, 1975-1976, *11*, 599-622.

Rickards, J. P. Interaction of position and conceptual level of adjunct questions on immediate and delayed retention of text. *Journal of Educational Psychology*, 1976, *68*, 210-217. (a)

Rickards, J. P. Type of verbatim question interspersed in text: A new look at the position effect. *Journal of Reading Behavior*, 1976, *8*, 37-45. (b)

Rickards, J. P. Stimulating high-level comprehension by interspersing questions in text passages. *Educational Technology*, 1976, *16*, 13-17. (c)

Rickards, J. P., Anderson, M. C., & McCormick, C. B. Processing effects of common-word and number questions inserted in reading materials. *The Journal of Educational Research*, 1975-1976, *69*, 274-277.

Rickards, J. P., & August, G. J. Generative underlining strategies in prose recall. *Journal of Educational Psychology*, 1975, *67*, 860-865.

Rickards, J. P., & DiVesta, F. J. Type and frequency of questions in processing textual material. *Journal of Educational Psychology*, 1974, *66*, 354-362.

Rivera, T. M. A reading program using tutorial assistance for selected educational opportunity program students. (Doctoral dissertation, University of California, Los Angeles, 1973.) *Dissertation Abstracts International*, 1974, *34*, 5600A. (University Microfilms No. 74-3984.)

Robin, A. L. Behavioral instruction in the college classroom. *Review of Educational Research*, 1976, *46*, 313-354.

Romine, S. Perceptions of an effective community college instructional climate. *Journal of Higher Education*, 1974, *45*, 415-429.

Rosenkranz, C. I. R. The efficacy of cloze procedure for estimating reading ability of students and readability of materials in adult fundamental education programs. (Doctoral dissertation, University of Wisconsin-Madison, 1975.) *Dissertation Abstracts International*, 1976, *36*, 4964A. (University Microfilms No. 75-23, 868.)

Ross, S. M., & DiVesta, F. J. Oral summary as a review strategy for enhancing recall of textual material. *Journal of Educational Psychology*, 1976, *68*, 689-695.

Rothkopf, E. Z. The concept of mathemagenic activities. *Review of Educational Research*, 1970, *40*, 325-336.

Rothkopf, E. Z. Structured text features and the control of processes in learning from written materials. In R. O. Freedle & J. B. Carroll (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D.C.: V. H. Winston and Sons, 1972.

Rothkopf, E. Z., & Billington, M. J. Indirect review and priming through questions. *Journal of Educational Psychology*, 1974, *66*, 669-679.

Rothkopf, E. Z., & Billington, M. J. A two-factor model of the effect of goal-descriptive directions on learning from text. *Journal of Educational Psychology*, 1975, *67*, 692-704. (a)

Rothkopf, E. Z., & Billington, M. J. Relevance and similarity of text elements to descriptions of learning goals. *Journal of Educational Psychology*, 1975, *67*, 745-750. (b)

Rothkopf, E. Z., & Kaplan, R. Exploration of the effect of density and specificity of instructional objectives on learning from text. *Journal of Educational Psychology*, 1972, *63*, 295-302.

Royer, J. M., & Cable, G. W. Facilitated learning in connected discourse. *Journal of Educational Psychology*, 1975, *67*, 116-123.

Royer, J. M., & Cable, G. W. Illustrations, analogies, and facilitative transfer in prose learning. *Journal of Educational Psychology*, 1976, *68*, 205-209.

Rubin, D. C. The effectiveness of context before, after, and around a missing word. *Perception and Psychophysics*, 1976, *19*, 214-215.

Russell, S. N. Error pattern relationships of developmental readers and functionally illiterate adults. (Doctoral dissertation, Oklahoma State University, 1973.) *Dissertation Abstracts International*, 1974, *34*, 6273A-6274A. (University Microfilms No. 74-8966.)

Rutkowski, K., & Domino, G. Interrelationship of study skills and personality variables in college students. *Journal of Educational Psychology*, 1975, *67*, 784-789.

Salomon, G. Heuristics for generation of aptitude x treatment interaction hypotheses. *Review of Educational Research*, 1972, *42*, 327-343.

Salup, B. J. An investigation of the cloze technique for measuirng the reading comprehension of college freshmen. (Doctoral dissertation, Temple University, 1974.) *Dissertation Abstracts International*, 1975, *36*, 3525A-3526A. (University Microfilms No. 75-28,139.)

Samuels, S. J., Begy, G., & Chen, C. C. Comparison of word recognition speed and strategies of less skilled and more highly skilled readers. *Reading Research Quarterly*, 1975-1976, *11*, 72-87.

Samuels, S. J., & Dahl, P. R. Establishing appropriate purpose for reading and its effect on flexibility of reading rate. *Journal of Educational Psychology*, 1975, *67*, 38-43.

Sanders, J. R. Retention effects of adjunct questions in written and aural discourse. *Journal of Educational Psychology*, 1973, *65*, 181-186.

Santa, C. M. Spelling patterns as units of word perception for adults and children reading on first, second and fifth grade reading levels. (Doctoral dissertation, Temple University, 1976.) *Dissertation Abstracts International*, 1976, *36*, 7963A. (University Microfilms No. 76-12,027.)

Sawyer, D. J. Intra-sentence grammatical constraints in readers' sampling of the visual display. *The Journal of Educational Research*, 1975-1976, *69*, 198-201.

Schallert, D. L. Improving memory for prose: The relationship between depth of processing and context. *Journal of Verbal Learning and Verbal Behavior*, 1976, *15*, 621-632.

Schnell, T. R. The effect of organizers on reading comprehension of community college freshmen. *Journal of Reading Behavior*, 1973, *5*, 169-176.

Schnell, T. R. The community college reading teacher: A profile. *Journal of Reading*, 1974, *18*, 8-11.

Schoenfeldt, L. F., Schoenfeldt, B. B., Acker, S. R., & Perlson, M. R. Content validity revisited: the development of a content-oriented test of industrial reading. *Journal of Applied Psychology*, 1976, *61*, 581-588.

Schumacher, G., Liebert, D., & Fass, W. Textual organization, advance organizers, and the retention of prose material, *Journal of Reading Behavior*, 1975, *7*, 173-180.

Schwartz, S. Encoding specificity and recognition memory for words. *Bulletin of the Psychonomic Society*, 1975, *6*, 279-281.

Schweitzer, J. C. Comparison of young subscribers and nonsubscribers to newspapers. *Journalism Quarterly*, 1976, *53*, 287-293.

Shavelson, R. J., Berliner, D. C., Ravitch, M. M. & Loeding, D. Effect of position and type of question on learning from prose material: Interaction of treatments with individual differences. *Journal of Educational Psychology*, 1974, *66*, 40-48.

Siegel, A. I., Lautman, M. R., & Burkett, J. R. Reading grade level adjustment and auditory supplementation as techniques for increasing textual comprehensibility. *Journal of Educational Psychology*, 1974, *66*, 895-902.

Sissors, J. Z. Do youthful college-educated readers prefer contemporary newspaper designs? *Journalism Quarterly*, 1974, *51*, 307-313.

Sloboda, J. A. The effect of item position on the liklihood of identification by inference in prose reading and music reading. *Canadian Journal of Psychology*, 1976, *30*, 228-237.

Slovak, P. A. The reading achievement levels of community college freshmen enrolled in a modular English course compared with the readability levels of books assigned in the course. (Doctoral dissertation, East Texas State University, 1975.) *Dissertation Abstracts International*, 1976, *36*, 5096A. (University Microfilms No. 76-4658.)

Smith, B. D. The effect of differential diagnosis upon reading improvement for high risk junior college students. (Doctoral dissertation, Georgia State University, School of Education, 1974.) *Dissertation Abstracts International*, 1975, *35*, 4995A-4996A. (University Microfilms No. 74-23,178.)

Smith, C. L. G. A comparative study of the readability of selected accounting textbooks. (Doctoral dissertation, University of Oklahoma, 1975.) *Dissertation Abstracts International*, 1975, *36*, 1990A. (University Microfilms No. 75-21,827.)

Smith, G. D., Enright, G., & Devirian, M. A national survey of learning and study skills programs. *Reading: Convention and inquiry*. National Reading Conference Yearbook, 1975, *24*, 67-73.

Smith, M. L., and Glass, G. V. Meta-analysis of phychotherapy outcome studies. *American Psychologist*, 1977, *32*, 752-760.

Snowman, J., & Cunningham, D. J. A comparison of pictorial and written adjunct aids in learning from text. *Journal of Educational Psychology*, 1975, *67*, 307-311.

Solis, J. K. E. Factors related to and predictive of residual reading gains for a junior college population. (Doctoral dissertation, New Mexico State University, 1973.) *Dissertation Abstracts International*, 1974, *34*, 5807A. (University Microfilms No. 74-7091.)

Somekh, D. E. The effect of embedded words in a brief visual display. *British Journal of Psychology*, 1976, *67*, 529-535.

Spears, C. D. A comparative analysis of an adult learning center and a formal adult education program with reference to selected variables. (Doctoral dissertation, University of Michigan, 1973.) *Dissertation Abstracts International*, 1974, *34*, 4680A. (University Microfilms No. 74-3728.)

Steingart, S. K. The effects of imagery and text structure on what is learned from reading a passage. (Doctoral dissertation, Cornell University, 1975.) *Dissertation Abstracts International*, 1975, *36*, 3528A. (University Microfilms No. 75-27,862.)

Stevens, G. L. An investigation of the relationship between adult attitudes toward reading and reading skills before and after a reading improvement course. (Doctoral dissertation, University of Maryland, 1974.) *Dissertation Abstracts International*, 1975, *36*, 1245A-1246A. (University Microfilms No. 75-17,907.)

Stovall, J. L. A study of attitudes of certain groups in Harris County toward remedial programs for all age groups in the Houston community college system. (Doctoral dissertation, University of Houston, 1975.) *Dissertation Abstracts International*, 1975, *36*, 2676A. (University Microfilms No. 75-23,949.)

Stumpf, A. D. Relations of reading ability and prefatory objectives to immediate achievement. (Doctoral dissertation, University of Missouri-Columbia, 1974.) *Dissertation Abstracts International*, 1975, *36*, 853A-854A. (University Microfilms No. 75-16,063.)

Sullivan, E. P. Internal reading flexibility patterns among Texas A & M University undergraduates. (Doctoral dissertation, Texas A & M University, 1974.) *Dissertation AbstrACTS International*, 1975, *35*, 5068A. (University Microfilms No. 75-2899.)

Tanenhaus, M. K., Carroll, J. M., & Bever, T. G. Sentence-picture verification models as theories of sentence comprehension: A critique of Carpenter and Just. *Psychological Review*, 1976, *83*, 310-317.

Thomas, G., Hartley, R. D., & Kincaid, J. P. Test-retest and interanalyst reliability of the automated readability index, Flesch reading ease score, and the Fog count. *Journal of Reading Behavior*, 1975, *7*, 149-154.

Thomas, L. E. Listening as an adjunct to remedial reading at the junior community college level. (Doctoral dissertation, North Carolina State University at Raleigh, 1974.) *Dissertation Abstracts International*, 1975, *36*, 1246A. (University Microfilms No. 75-15,169.)

Tillman, C. E., Millott, R. F., & Larsen, J. J. Using personality test data in individualized reading instruction: Brief reports of five studies of personality and reading. *Interaction: research and practice in college-adult reading*. National Reading Conference Yearbook, 1974, *23*, 125-130.

Tinkle, J. L. A study of two teaching schedules of teaching functionally illiterate adults to read at the Indiana State Penal Farm. (Doctoral dissertation, Indiana University, 1973.) *Dissertation Abstracts International*, 1974, *34*, 4682A. (University Microfilms No. 74-2714.)

Trachtman, J. P. Cognitive and motivational variables as predictors of academic performance among disadvantaged college students. *Journal of Counseling Psychology*, 1975, *22*, 324-328.

Travers, J. R. Forced serial processing of words and letter strings: A reexamination. *Perception and Psychophysics*, 1975, *18*, 447-452.

Treppa, J. A. Personality characteristics of successful college students who enroll in study skills courses. *Journal of Counseling Psychology*, 1973, *20*, 546-552.

Underwood, G. Semantic interference from unattended printed words. *British Journal of Psychology*, 1976, *67*, 327-338.

Voss, J. F. Acquisition and nonspecific transfer effects in prose learning as a function of question form. *Journal of Educational Psychology*, 1974, *66*, 736-740.

Wark, D. M., & Walch, M. Heart rate of high and low test-panic readers. *Interaction:Research and practice in college-adult reading.* National Reading Conference Yearbook, 1974, *23*, 209-213.

Warren, M. O. N. A survey of reading and study skills programs in selected Rocky Mountain colleges and universities. (Doctoral dissertation, University of Wyoming, 1974.) *Dissertation Abstracts International*, 1975, *36*, 720A. (University Microfilms No. 75-12,846.)

Weener, P. Note taking and student verbalization as instrumental learning activities. *Instructional Science*, 1974, *3*, 51-74.

Weiss, L. S. A compensatory reading program for the community junior college. (Doctoral dissertation, Florida Atlantic University, 1974.) *Dissertation Abstracts International*, 1974, *35*, 1436A-1437A. (University Microfilms No. 74-20,597.)

Whittaker, J. W. P. Effects of the application of linguistics on reading comprehension of black freshmen students at Texas Southern University. (Doctoral dissertation, University of Houston, 1974.) *Dissertation Abstracts International*, 1974, *35*, 1886A-1887A. (University Microfilms No. 74-21,402.)

Wickelgren, W. A. Age and storage dynamics in continuous recognition memory. *Developmental Psychology*, 1975, *11*, 165-169.

Wiggins, N. D. Cumulative reading deficits of freshman college students. (Doctoral dissertation, Oklahoma State University, 1974.) *Dissertation Abstracts International*, 1975, *35*, 6535A. (University Microfilms No. 75-8915.)

Williams, L. L. Use of galvanic skin response, heart rate, respiration rate, Betts's criteria, and self-rating scales to compare college students' degrees of frusteration on informal reading inventories. (Doctoral dissertation, University of Pittsburgh, 1975.) *Dissertation Abstracts International*, 1976, *36*, 8015A-8016A. (University Microfilms No. 76-14,184.)

Wilson, R. C., & Einbecker, P. G. Does reading ability predict college performance? *Journal of Reading*, 1974, *18*, 234-237.

Winter, K. K. An investigation of the influences of context on sentence reading. (Doctoral dissertation, Cornell University, 1974.) *Dissertation Abstracts International*, 1974, *35*, 734A. (University Microfilms No. 74-17,693.)

Wisher, R. A. The effects of syntactic expectations during reading. *Journal of Educational Psychology*, 1976, *68*, 597-602.

Wolfe, R. F. An examination of the effects of teaching a reading vocabulary upon writing vocabulary in student compositions. (Doctoral dissertation, University of Maryland, 1975.) *Dissertation Abstracts International*, 1975, *36*, 3337A. (University Microfilms No. 75-28,946.)

Wood, D. S. The relationship between rate controlled speech and reading comprehension skills of postsecondary technical students. (Doctoral dissertation, Memphis State University, 1975.) *Dissertation Abstracts International*, 1976, *36*, 5797A-5798A. (University Microfilms No. 76-4604.

Yekovich, F. R., & Kulhavy, R. W. Structural and contextual effects in the organization of prose. *Journal of Educational Psychology*, 1976, *68*, 626-635.

Zilman, D., & Cantor, J. R. Induction of curiosity via rhetorical questions and its effects on the learning of factual materials. *British Journal of Educational Psychology*, 1973, *43*, 172-180.

PAPERS PRESENTED AT THE 1977 MEETING OF THE NATIONAL READING CONFERENCE IN NEW ORLEANS

For the convenience of NRC members who may wish to write directly for papers, we are listing all papers presented at the New Orleans conference. We have systematically excluded program chairpersons, discussants, respondents, etc. Also we have tried to make all changes in paper authorship and titles; however it is likely that there still exist some discrepancies.

1.1 George McConkie, Cornell University. Visual Perception.
1.2 Gregory Pearson, Cornell University. The Acquisition of Knowledge from Text.
2.1 Dawn C. Mayhew, University of Georgia. Kindergarten Children's Orthographic Representation of Selected Phonemes.
2.2 Jerome Niles and Barbara Taylor, Virginia Polytechnic Institute and State University. The Development of Orthographic Sensitivity During the School Year by Primary Grade Children.
2.3 Nancy E. Taylor, Catholic University. Children's Awareness of Word Boundaries as a Cueing Convention in Written Language.
2.4 Josephine Goldsmith, Rutgers University. Word Boundary Decisions as a Function of Interest for Developmental and Retarded Readers.
3.1 Thomas Bean, Western Illinois University. Influence of a Structured Overview on Comprehension and Oral Reading Miscues of Selected College Students.
3.2 Richard Vacca, University of Connecticut. The Effect of Expanded Directions and Adjunct Aid Placement on Reading Comprehension of Social Studies Text Material.
3.3 James Riley, University of North Carolina. The Effect of Three Level Reading Guides on Word Problem Comprehension.
3.4 Anne Wolf, Oconomowoc Junior High School, Wisconsin. Effects of Sustained Silent Reading and Reading Skills Instruction on Secondary Student's Reading Attitudes and Achievement.
4.1 Chester Tillman, University of Florida. Reading Textbooks - SQ3R and All That.
4.2 Janet Larsen, University of Florida. Preparation for Exams: Diverse Approaches.
4.3 Alan Kirby and A. Garr Cranney, University of Florida. Time Use of College Students.
5.1 Richard Barron and Harold Cafone, Oakland University. Process/Product Confusion Among Disabled Readers.
5.2 Noel Bowling, West Virginia College of Graduate Studies and James Laffey, University of Virginia. Survey of Effective Teaching Practices.
5.3 Robert Blanc, University of Missouri, Kansas City. Intuition Training as a Component of Reading Programming.
5.4 Marc Glassman, Memorial University of Newfoundland. Comparison of Traditional and CBTE Introductory Reading Methodology Courses.
6.1 Carl Frederiksen, National Institute of Education. Discourse Analysis: Research and Implications.
6.2 Robert Tierney, C. Bridge, M. Cera, University of Arizona. Discourse Processing Operations of Children.
6.3 C. Bridge, R. Tierney, and M. Cera. Inferential Operations of Children Involved in Discourse Processing.
6.4 Nancy Marshall, University of Virginia. Inferred and Generated Information in Text Recalls of Good and Poor Readers.
7.1 Mark Aulls, McGill University. The Influence of Levels of Reading Fluency on Comprehension and Decoding for Good and Poor Readers in the Primary Grades.
7.2 James Flood and Diane Lapp, Boston University. The Relationship Between Inferential Reading Comprehension and Questioning Strategies.
7.3 James Laffey and Joseph Muia, James Madison University. Comprehension Strategies of Fourth, Fifth, and Sixth Graders.
7.4 Stanley Wanat, California State University, Fullerton. Semantic Factors as Mobilizers in Early Reading: A Reading Diary Study.
8.1 Anthony Manzo and John Sherk, University of Missouri, Kansas City. Reading and Languaging in the Content Areas.
8.2 James Hoffman, University of Missouri, Kansas City. An Investigation of the Intra-Act Procedure Within the Context of Languaging in the Content Area.
8.3 Marilyn Eanet, University of Missouri, Kansas City. An Investigation of the REAP/Study Procedure: Its Rationale and Efficacy.
8.4 Jane Meeks, University of Missouri, Kansas City. An Investigation Into the Effects of Imbedded Aids.
8.5 Martha Haggard, University of Missouri, Kansas City. Creative Thinking - Reading Activities (CT-RA) for Disabled Readers - An Alternative Approach.
9.1 Edward Summers, Anne Forester, and Sharon Jeroski, University of British Columbia. Review of Research on College and Adult Reading, 1975-1976.

9.2 Edward Summers and David Barnett, University of British Columbia. Using Phenomenalistic Data to Characterize Reading Research.

9.3 Hal Seaton and M. Jean Greenlaw, University of Georgia. Publishing in Reading: Policies, Trends and Problems.

10.1 Allen Berger and James Peebles. A Review of ERIC Documents About Reading Disability and Learning Disability.

10.2 Ned Ratekin and Catherine Hatcher, University of Northern Iowa. The Effect of the Structure of Materials on the Development of Reading Skills in Disabled Readers.

10.3 James McLeskey, Georgia State University. The Acquisition of a Conditional Discrimination Learning Set by Reading Disabled and Normal Children.

10.4 Bernard Karmel, University of Connecticut. Neurometrics Related to Reading Dysfunction.

10.5 Patricia Duncan, Virginia Commonwealth University. Cognitive Sequencing of Learning Disabled Students Following Listening and Reading Tasks.

11.1 Glenn Kleiman, University of Illinois at Champaign-Urbana. Perspectives on Analyzing Written and Oral Language Comprehension.

11.2 Diane Schallert, University of Illinois at Champaign-Urbana. Empirical Investigations of the Effects of Differences Between Oral and Written Language on Comprehension.

11.3 Ann Rubin. Context-Shifting: Its Role in Transition from Oral Language Comprehension to Reading.

11.4 Erica McClure. Relationship Between Non-Standard English Dialects and Standard Written English.

12.1 Ray McDermott, Rockefeller University. The Cultural Context of Reading.

13.1 Mary Dupuis, Pennsylvania State University. Content Area Reading Project: Model and Evaluation Design.

13.2 Joyce Lee. Presentation of Results from the Content Reading Project: Changes in Knowledge and Attitudes.

13.3 Eunice Askov. Presentation of Results from the Content Area Reading Project: Factors Related to Change.

14.1 Roger Quealy, University of Wisconsin-Eau Claire. Video Tape Based Reading Improvement Program for a Black Junior College.

14.2 Alice Grant. The Black Junior College Student.

14.3 June Gilstad. A Black Junior College Reading Program.

15.1 Ruth Meyers, New York University. Information Processing as a Tool for Selecting Instructional Strategies.

15.2 Margaret Smith-Burke, New York University. Psycholinguistics as a Base for Understanding Reader/Text Interaction.

16.1 Laurence Iannaccone, University of California, Santa Barbara. Education Policy Changes and the Politics of Reading.

17.1 George Mason, University of Georgia. Word Boundaries: An International Assessment at the Completion of a Year of Schooling in Different Countries: Introduction.

17.2 Sandra Mallula. Finnish.

17.3 Jane Higdon. Arabic (Morocco).

17.4 John Mize. Farsee (Iran).

17.5 Carol Webb. Afrikaans (South Africa).

17.6 George Mason. English (Southeastern USA).

17.7 Ann Hall. French (South Louisiana, USA).

17.8 Nelly Hecker. Hebrew (Israel).

17.9 Margaret Carswell. Italian (Italy).

17.10 Ann Elam and/or Linda Aruthur. Japanese (Japan).

17.11 Nelly Hecker. Spanish (Colombia).

18.1 Emery Bliesmer and Emily Walcher, Pennsylvania State University. Preservice Preparation of Teachers for Teaching Reading at Various Levels.

18.2 Ira Aaron and Bob Jerrolds, University of Georgia. Middle School Level.

18.3 Dan Dolan, University of California, Riverside. High School Level.

18.4 Alton Raygor and Bill Browning, University of Minnesota. College Level.

19.1 Susan Brown and Leon Williamson, New Mexico State University. The Value of the Reading Miscue Inventory for Evaluating College Freshmen Reading Performance.

19.2 Tomi Berney, Rutgers University. Relationships Between Questionnaire Responses and Subsequent Standardized Reading Test Performance.

19.3 Lynda Craft and Ronald Noland, Madison County Schools/Auburn University, Auburn, Alabama. The Application and Analysis of the Granowsky, Botel, and Dawkins Syntactic Complexity Formula as Applied to Social Science Textbooks to Determine Readability.

19.4 Charles Peters, Oakland Schools, Pontiac, Michigan. Assessing Reading Ability at the Secondary Level Through the Utilization of a Cognitive Inventory.

20.1 Ladessa Yuthas, Metropolitan State College, Denver. The Effect of a Combined Reading-Writing Course on Achievement of Probation Students.

20.2 Cathy Stallard and Caesarea Abartis, Southern Illinois University at Carbondale. The Effect of Formal Instruction in Writing Composition or Reading Methodology Upon the Reading Skills of College Students.

20.3 Carol Santa and Lindsay Abrams, Rutgers University. The Effect of Note Taking on the Retention of Prose Material.

20.4 Neil Kaufman, Case-Western Reserve University. A Hierarchical-Mnemonic Approach to the Learning and Retention of Prose.

21.1 J. Bleakley and E. Johnsen, University of Kansas. The Comprehension of Prose: Problems in Measuring Learning Outcomes.

21.2 Walter Prentice and Joe Peterson, University of North Dakota. Beyond Passage Dependency: A Closer Look at What Tests of Reading Comprehension Measure.

21.3 Nancy Marshall, University of Virginia. Qualitative Analysis of Uniquely Recalled Items in Subjects' Free Recalls: A Differential Description of Good and Poor Readers' Inferences.

21.4 Donald Cunningham, Indiana University. The Retention of Connected Discourse, II.

22.1 Elton Stetson, University of Houston. One Hundred Nineteen Phonograms: An Analysis of Pronunciability.

22.2 George McNinch and Wallace Miller, University of Southern Mississippi. Experiments in Word Learning.

22.3 Jeannie Steele and James Laffey, James Madison University. Word Identification Strategies.

22.4 Darla Cohen, Center for Innovation in Teaching the Handicapped, Bloomington, Indiana. A Comparison of Reader Perceptions of Decoding Strategies With Strategies Readers Actually Used.

23.1 Edward Paradis, University of Wyoming and Joe Peterson, University of North Dakota. Effects of Order of Administration of Subtests on Standardized Reading Test Scores.

23.2 Carol Davis, Purdue University. The Effectiveness of Informal Assessment Questions Constructed by Secondary Teachers.

23.3 Jill Olshavsky, Indiana University. How Readers Fail to Cope With Material of Increasing Difficulty.

23.4 Jocelyn Chen, IBM Corporation, and Richard Bloomer, University of Connecticut. Learning Profiles of Good and Poor Readers.

24.1 Judith Irwin and Paul Blohm, University of Wisconsin-Madison. The Relationships Between and Among the Self-Reported Vividness of Mental Imagery, Reading Comprehension Ability, and Reading Attitude of College Readers.

24.2 Sandra Burkett, Mississippi State University. A Comparison of Successful and Unsuccessful Economic and Educational Disadvantaged College Students on Selective Intellectual and Non-Intellectual Factors.

24.3 Elois Skeen and Richard Scott, State University of New York at Buffalo. The Relationship of Reading Achievers to the Cognitive Style of College Learning Center Students.

24.4 Marilyn Fairbanks and Betsy Hobbs, West Virginia University. Reading Achievement Scores and Spelling Patterns of College Students.

25.1 Carol Greenfield, University of Northern Iowa. The Reading Readiness of Future Teachers.

25.2 Dianne Hunter, University of Missouri, Columbia. First Grade Teachers' Instructional Preferences for Reading: Implications for Research.

25.3 Karl Koenke, University of Illinois. A Minimal Professional Library in Elementary Education: A Survey.

25.4 Glenn Freedman, University of Houston-Clear Lake City. Sociolinguistic Boundary Effects on Reading Instruction.

25.5 Timothy Blair, University of Florida. Research on Teacher Effectiveness in Reading: Past, Present, and Future.

26.1 P. David Pearson, University of Minnesota. Semantic Categories and Reading.

26.2 Karen Hanson, Normandale, and Robert Schreiner, University of Minnesota. The Relationship Between Reading Ability and Semantic Verification Tasks.

26.3 John Childrey and Michael Kamil, Purdue University. A Preliminary Examination of the Relationship Between Category Structure and Reading.

26.4 Michael Kamil and Raymond Hanson, Purdue University. Perceptual vs. Semantic Information Processing.

27.1 David Elkind, University of Rochester. Stages of Reading Development.

27.2 Irene Athey, University of Rochester. Response.

28.1 Marie Bolchazy, State University of New York at Albany. Teaching Reading in the Content Areas: A Pilot Study.

28.2 Effie Adams, Jarvis Christian College. Professional Reading Teachers in Bureaucratic Organizations.

28.3 Joseph Mahony and Fred Linder, Virginia Commonwealth University. A Survey of Secondary Reading Programs in Virginia.

28.4 Sheilah Allen, University of British Columbia. A Model for Teacher Education in Secondary Reading.

29.1 Anne Eisenberg, New York City Community College. Syntactic and Semantic Characteristics of the Language of Science in the Teaching of Reading Using Science - Content Materials.

29.2 Elliott Mellichamp, Tennessee State University. Investigating the Nature of Main Ideas of Sentences to Improve Comprehension of College Textbooks.

29.3 Martha Maxwell, University of California, Berkeley. Applying Information Processing Theory and Research to Basic Reading Materials.

30.1 Gary Shaffer and Ann Marie Leonard, James Madison University. Television and Reading: A Summary of the Literature.

30.2 Mary Lorton, Aquinas College and Christopher Kutuk, University of Missouri, Kansas city. Sociological and Social Process Variables In Developmental and Delayed Readers.

30.3 Reta Hicks and Eula Monroe, Western Kentucky University. A Comparative Analysis Among the Variables of Intelligence, Sex, and Gain Made in Comprehension of Sixth Year Students When Taught Global or Specific Comprehension Skills.

30.4 Mary Creamer, West Georgia College. The Influence of Race, Sex, and Broken Homes on the Reading Achievement of Sixth Graders.

31.1 Priscilla Drum, University of California, Santa Barbara. Inferences as Text Entailment.

31.2 Frank Flynn, Boston University. Effects of Contradiction on Inference.

31.3 Carol Dixon, University of California, Santa Barbara. An N of 1, Case Study: Language Experience and Inference.

31.4 Roger Lantaff, University of California, Santa Barbara. The Effects of Length, Depth, and Height on Text Recall.

31.5 James Flood, Boston University. Cumulative Effects of Information on Inference-making.

32.1 Joel Levin, University of Wisconsin, and Alan Lesgold, University of Pittsburgh. Do Pictures Improve Children's Prose Learning? An Examination of the Evidence.

32.2 Brenda Kolker, University of Tennessee. Effects of Imagery on Number of Trials to Criterion Learning of First Grade Students.

32.3 George Lamb and John Towner, Western Washington State College. Phonological Coding in Letter Discrimination.

32.4 Elizabeth Tucker, University of Virginia. Word Sort Tasks in Relation to Word Knowness.

33.1 Earl Rankin, University of Kentucky. The Use of the Cloze Procedure as a Research Tool in the Study of Language.

33.2 Martin Kling and Donald Pratt, Rutgers University. Readability: A Re-examination and a Study Indicating New Directions for Further Research.

33.3 Michael McKenna, Wichita State University. Computer Simulation of Cloze Performance.

33.4 Richard Allington, State University of New York at Albany. Prediction Processes in Good and Poor Readers.

33.5 Jerry Johns, Northern Illinois University. An Investigation Using the Cloze Procedure as a Teaching Machine.

34.1 Tom Duffy, Navy Personnel R and I Center, San Diego. Semantic and Syntactic Factors in Reading Test Performances.

34.2 E. Jongsma, Southern Methodist University. The Effects of Instruction in Testwiseness in a College Reading Improvement Course.

34.3 Eunice Schmidt, Seattle Pacific University. What Makes Reading Difficult? The Complexity of Structures.

34.4 Lyndell Grey, University of Georgia. Performances of High, Moderate, and Low Reading Subjects on a Read and Record Task.

34.5 M. Jane Greenewald, University of Wisconsin-LaCrosse. Direct and Indirect Measures of Reader-Response to Three Levels of Language Usage.

35.1 George Canney, University of Illinois. Criteria for Selecting More Effective Worksheet Exercises in Reading.

35.2 Homer Coker, West Georgia College. Observed Patterns of Teacher-Pupil Classroom Behavior as Predictors of Pupil Growth in Reading.

35.3 Robert Aaron, University of Georgia. Time-Use: A Variable in Teacher Effectiveness.

35.4 Joanne Olson, University of Houston. Graduate Students' Perceptions of Selected Components of a Competency Based Reading Course.

35.5 Shirley Merlin, James Madison University. The Effect of Modular-Non Modular Instruction on Student Competency in a Preservice Education Program.

36.1 Joanna Williams, Columbia University. Summative data on the ABD's of Reading.

36.2 Tom Donlon, Educational Testing Service. Response Validation in Reading Comprehension.

36.3 Ann Bussis, Edward Chittenden, Marianne Amarel, Educational Testing Service. An Approach to the Study of Beginning Reading: Longitudinal Case Studies.

36.4 Roy Freedle, Educational Testing Service. Reading Frames, Culture Frames, and Language Frames: Their Mutual Interactions.

37.1 Carol Hopkins, Purdue University. The Prediction of Third-Grade Reading Achievement from Selected Measures of First-Grade Oral Language.

37.2 Alden Moe, Purdue University. Predicting First-Grade Reading Achievement from Selected Measures of Oral Language Performance.

37.3 Robert Rush and Alden Moe, Purdue University. Correspondence Between Syntactical Features of Primary Basal Reader Language and the Oral Language of Beginning First-Grade Children.

38.1 Robert Wisher and Ed Aiken, Navy Personnel Research and Development Center, San Diego. Information Overload During Reading.

38.2 Nancy Nelson and John Geyer, Rutgers University. Immediate Memory for Sentences of Fast and Slower Readers as a Function of Rate of Presentation.

38.3 Fred Ohnmacht and James Fleming, State University of New York at Albany. Cognitive Processing Differences of Good and Poor Readers.

38.4 Connie Bridge, University of Arizona. Inferential Operations of Children Involved in Discourse Processing.

39.1 Floyd Sucher, Brigham Young University. Reading Interest and Study Habits of University Students.

39.2 Daryl Dickson and Rowena Wilhelm, University of Michigan. The Effect of Teacher-Centered and Learner-Centered Delivery Systems on Learner Performance.

39.3 Vincent Orlando and Kristina Hayward, Pennsylvania State University. Comparison of the Effectiveness of Three Study Techniques for College Students.

39.4 Gordon Cox, University of California, Berkeley. Differential Effects of Instruction Stressing Cognitive Control Strategies vs. Diagnostic Prescriptive Instruction on Reading Rate Flexibility and Comprehension of College Students.

39.5 Joseph Ilika and Bonnie Longnion, Texas A&M University. A Comparison of a Five and One Half and an Eleven Week Reading Improvement Course for Government Employees.

40.1 Wayne Otto and Betty Harper, University of Wisconsin. Constructing a Test of Teachers' Knowledge of Reading.

40.2 Gerald Duffy, Michigan State University; Wayne Otto, and S. Jay Samuels. A Study of Teacher Conceptions of Reading.

40.3 Yvonne Steinruck, Southern Illinois University. Teaching Reading in the Elementary School: Results of a Course Evaluation by Student Teachers, Cooperating Teachers, and University Supervisors.

40.4 Jo Anne Vacca, Boston University. Evaluating Staff Development Personnel.

40.5 Harry Singer, University of California, Riverside. Teachers' Attitudes Towards Teaching Students to Read vs. Learn from Text.

41.1 Laurence Frase, National Institute of Education. The Psychology of Document Design.

41.2 S. Jay Samuels, University of Minnesota. Word recognition: Letter-by-Letter or Chunk - Resolved?

41.3 Patricia Cunningham, Alamance County Schools and James Cunningham, University of North Carolina, Chapel Hill. Investigating the "Print to Meaning" Hypothesis.

41.4 Leo McCusker, University of Texas at Austin; Michael Cosky, St. Olaf College; and Philip Gough, University of Texas at Austin. Phonetic Interference in Proofreading: Evidence for a Phonological Stage in Reading.

42.1 Thomas Gee, University of Houston at Clear Lake City. Preparation of a Group Informal Reading Inventory for the Social Studies.

42.2 Jo Ellen Oliver and Robert Palmatier, University of Georgia. Validation of a Criterion-Referenced Test Measuring Thinking Prerequisites to Comprehension.

42.3 Thomas Hatcher and Mary Yeazell, West Virginia University. Criterion Referenced Management Programs and Piagetian Theory.

42.4 Thomas Estes, University of Virginia. The Relationship of Learning Environment, Teacher Characteristics, and Student Attitudes Toward Math, Science, and Reading.

42.5 Ruth Kurth and Patricia Moseley, North Texas State University. Using Research on Readability to Enhance Middle School Curriculum Materials.

43.1 Albert Kingston, University of Georgia. Introduction to Enhancing the Lives of Aged People Through Reading.

43.2 Molly Wilson, University of Georgia. Future Shock and the Elderly: Is Reading a Cure or Part of the Problem?

43.3 Terry Lovelace, Louisiana State University. Problems in the Measurement of Intellectual and Linguistic Capabilities and Achievement, Including Reading, Among the Aged.

43.4 Ann Marie Bernazza Haase, Kenneth Smith, and Helen Renthal, University of Arizona. Agism in Children's Literature.

43.5 Richard Robinson, University of Missouri, Columbia. The Usefulness of Readability Formulas in Comparison to the Cloze Procedure for the Selection of Appropriate Reading Material With the Aged.

44.1 Edward Fry, Rutgers University. Partial Validation of the Kernel Distance Theory.

44.2 Mary Herndon, San Diego State University. An Approach Toward the Production of Syntactically Smooth Reduced Text.

44.3 M. Jean Greenlaw, Hal Seaton and Carol Fisher, University of Georgia. A Study of Vocabulary and Syntactic Complexity in Developmental First Grade Readers.

44.4 Edward Summers and Joyce Matheson, University of British Columbia. Generating Core Reading Vocabularies from Ranked Frequency Distributions.

44.5 Elizabeth Sakiey, Glassboro State College. Graphemic Syllables: Unweighted and Weighted Tests.

45.1 Joseph Vaughn and Keith Meredith, University of Arizona. Reliability of the Cloze Procedure as Assessments of Various Language Elements.

45.2 Keith Meredith and Joseph Vaughan, University of Arizona. Stability of Cloze Scores Across Varying Deletion Patterns.

45.3 Kenneth Smith, University of Arizona. Criterion Referenced Testing of Decoding and/or Comprehension Using Cloze Procedure.

45.4 Adela Stewart, University of Arizona. Use of Cloze Procedures to Establish Readability Levels in Spanish.

45.5 James Cunningham, University of North Carolina, and Robert Tierney, University of Arizona. A Comparative Analysis of Cloze and Modified Cloze Procedures.

45.6 Ann Marie Bernazza Haase and John Bradley. Testing in the Content Areas Using Cloze and Maize Procedures.

45.7 John Bradley, University of Arizona. The Reliability of the Maize Procedure.

46.1 Allan Neilsen, University of Calgary and P. David Pearson, University of Minnesota. The Role of Macrostructures and Relational Markers in Comprehending Familiar and Unfamiliar Written Discourse.

46.2 Peter Mosenthal, State University of New York at Albany. Children's Acquisition of Discourse Structures in Three Modes of Comprehension.

46.3 Stephen Kucer, Indiana University. Cognitive Scripts and Reading Comprehension.

46.4 William Smith, Indiana University. Evidence of Information Processing After the Reading Act is Completed.

46.5 Laura Chodos, State University of New York at Albany. Children's Strategies in Comprehending Story Structures.

47.1 Margaret Smith-Burke, Daniel Eagleeye, and Patricia Gingrich, New York University. Review of Current Research on the Cloze Procedure as a Measure of Comprehension and Differential Effects of Prior Context on Cloze Performance in High School Readers.

47.2 Steven Kidder, New York State Education Department. Validation of the Multiple-Choice Exercises of the Test Development Notebook.

47.3 Michael Ravitch, University of Rochester. Relationship of Cloze Performance to Theories of Language and Comprehension.

47.4 Linda Evers and Earl Rankin, University of Kentucky. The Effect of Pretest Orientation Upon Cloze Test Performance.

48.1 Helen Covington and Lee Mountain, University of Houston. The Competencies That Junior College Chairpersons Expect from Their Course Graduates.

48.2 Deanna Martin, University of Missouri, Kansas City. Is Maximal Reading Achievement Optimal Reading Achievement?

48.3 Jeweleane Whittaker, Texas Southern University. The Effect of Competency-Based Reading and Study Skills Courses at the College Level.

48.4 Robert Pavlik, University of Northern Colorado. Improving the Reading-Writing Instructional Practices of University Faculty.

49.1 John Readence, Kansas State University. Effects of Impulsivity and Reflectivity and Type of Phonics Instruction on Reading Achievement.

49.2 Leo Campbell, Rutgers University. Cardiac Changes as a Result of Reader Passage Difficulty.

49.3 Jerome Harste, Indiana University. Content Area Reading and the Elementary School Child.

49.4 John Stansell, Texas A&M University. The Oral Reading and Retelling of Narrative and Expository Materials Selected by Ninth Graders.

49.5 Roger DeSanti, University of New Orleans. An Investigation of the Reading Strategies of Four Proficient Readers, 60 Years of Age or Older.

50.1 R. Ann Zinck, University of Georgia. An Investigation of Language Cues Utilized During Oral and Silent Reading.

50.2 Susan Pflaum and Tanis Bryan, University of Illinois at Chicago Circle. Methodology in Collection of Oral Reading Data.

50.3 Lyndon Searfoss, Arizona State University. A Comparison of Two Methods of Analyzing the Oral Reading Behaviors of Selected Remedial Readers.

50.4 Sally Kaminsky, College of Staten Island of the City of New York. Analysis of Language Development of Bilingual First Grade Children and Their Acquisiton of Reading.

50.5 Christopher Ramig, Georgia State University. Teacher Inferences of the Characteristics of Non-Standard Speaking Readers.

51.1 Sean Walmsley, State University of New York at Albany. Understanding of Logical Connectives in Sentence Memory Tasks.

51.2 Diana Lagotic, University of Florida. The Degree of Intra-Sentence Constraints of Four Types of Sentences as Measured by the Eye-Voice Span of Graduate Student Readers.

51.3 Patrick Finn, State University of New York at Buffalo. A Model of Reading Based on Case Grammar and Semantic Transfer Features.

51.4 Kathleen Gormley and Ann Franzen, State University of New York at Albany. Influence of Expectation on the Comprehension of Passive Sentences by Deaf Readers.

51.5 Charlene Andolina, University of Pittsburgh. Syntactic Maturity and Vocabulary Richness of Learning Disabled Children.

52.1 Ruth Hervatin, Case-Western Reserve University. Kindergarten Teachers' Attitudes Toward Factors Influencing Beginning Reading Achievement.

52.2 Penny Nielsen, Nicholls State University. Attitude Toward Reading and Reading-Related Concepts Among Elementary Students.

52.3 Anne Simpson, Texas A&M University. A Study of the Reading Interests of Early Adolescents and the Implications for Reading Programs.

52.4 Lana McWilliams, Memphis State University. Today's Illiterates and Tomorrow's Delinquents: Training Experiment on Locus of Control.

53.1 Mervin Lynch and Linda Lenz, Northeastern University. Levels of Reading Achievement: Their Impact on Self-Concept Related Frustration.

53.2 John Dixon, Texas Southern University. Effects of Group Reading Therapy on the Level of Reading and Personal Growth of Black Adolescents.

53.3 James Jackson, Southern Illinois University. The Teacher: A Necessary Element in the Use of Prereading and Concurrent Reading Activities: Relating New Learning to the Learners's Existing Knowledge.

53.4 Betty Yarborough, Old Dominion University. A Study of the Relative Effectiveness of Six Years of Non-Graded, Innovative Individualized Instruction vs. Six Years of Traditional Graded Instruction.

54.1 Ellen West, Alachua County Adult Education, Gainesville, Florida. The Use of An Unobtrusive Screening Device to Approximate Adult Reading Levels.

54.2 Philip Lanasa and Sue Sheridan, University of Houston at Clear Lake City. Development of a Criterion-Referenced Test and Instructional Program for the Trainable Mentally Retarded Adolescent.

54.3 Irwin Kirsch, International Reading Association. Concept and Measurement of Functional Literacy.

54.4 Mary Anne Hall, Georgia State University. A Survey of Adult Illiterates' Experiences with and Views of Reading.

54.5 Jane Higdon, University of Georgia. Training Practices in Adult Literacy Programs Using Volunteer Tutors.

55.1 Judy Mikulecky and Ken Dulin, University of Wisconsin-Madison. Plausibility Judgments of Arguments for and Against Reading as Indirect Indicators of Adult Reading Attitudes.

55.2 Ken Dulin, University of Wisconsin-Madison. The Indirect Measurement of Adult Readers' Attitudes Toward Books and Reading via a Projective Technique.

55.3 Patricia Anders, University of Arizona, Tucson. Adult Reading Habits and Self-Perceptions of Reading Ability in a Multi-Ethnic Community.